Лабораторная работа №4
Проверка статистических гипотез

Критерий Стьюдента

Теория

Напомним суть проблемы.

В случае одновыборочного теста мы будем исходить из того, что выборка X_1, X_2, ..., X_n распределена нормально. Необходимо оценить параметр нормального распределения μ — его математическое ожидание, а именно, проверить гипотезу $H_0 : \mu = \texttt{mu}$. В случае двухвыборочного теста мы имеем две нормально распределенные выборки X_1, X_2, ..., X_n и Y_1, Y_2, ..., Y_m. Необходимо проверить гипотезу о равенстве математических ожиданий этих распределений, а именно, $H_0 : \mu_x = \mu_y$ или $H_0 : \texttt{mu} = \mu_x - \mu_y = 0$

Синтаксис команды (разложен на строки для удобства описания):

```
t.test(
       x,
       y = NULL,
       alternative=c("two.sided","less","greater"),
       mu = 0,
       paired = FALSE,
       var.equal = FALSE,
       conf.level = 0.95,
   ...)
```

`x` вектор-выборка, распределённая по нормальному закону; если `y` отсутствует, то мы имеем дело с одновыборочным критерием Стьюдента; в противном случае имеем двухвыборочный критерий Стьюдента.

`y` вектор-выборка, распределённая по нормальному закону; при её наличии имеем двухвыборочный критерий Стьюдента.

`alternative=c("two.sided","less","greater")` вид альтернативной гипотезы:
двусторонняя `"two.sided"` $H_1 : \mu_x \neq \texttt{mu}$ или $H_1 : \mu_x \neq \mu_y$,
односторонняя `"less"` $H_1 : \mu_x < \texttt{mu}$ или $H_1 : \mu_x < \mu_y$,
односторонняя `"greater"` $H_1 : \mu_x > \texttt{mu}$ или $H_1 : \mu_x > \mu_y$;
мы можем выбрать один из вариантов или не выбирать никакой (в последнем случае это эквивалентно выбору варианта `"two.sided"`.

`mu` проверяемое значение математического ожидания или разность между математическими ожиданиями двух выборок.

`paired` логический параметр; если равен `TRUE`, то имеем связанные выборки `x` и `y`; в противном случае имеем независимые выборки `x` и `y`.

`var.equal` логический параметр; если равен `TRUE`, то подразумевается, что истинные дисперсии выборок неизвестны, но равны, в противном случае — неизвестны и не равны. Заметим, что до сих пор не найдено точное решение задачи сравнения средних двух выборок в том случае, когда дисперсии выборок неизвестны и не равны (проблема Беренса-Фишера). Поэтому на практике используются различные приближенные решения, два из которых — критерий Крамера-Уэлча и критерий Саттервайта. Именно они и применяются в R. В случае одновыборочного теста этот параметр отдельно не указывается. Судя по всему, полагается, что нам дисперсия не известна.

`conf.level` уровень достоверности, равный $1-\alpha$, где α — уровень значимости (по умолчанию полагаемый 0.05).

Задания

Задание 1. Проверьте гипотезы $H_0 : \mu = 8$ и $H_0 : \mu = 6$ при всех возможных видах конкурирующих гипотез, используя выборку на листе `Стьюдент1` из файла `данные.xlsx`. Напишите скрипт, который это выполняет.

Задание 2. Проверьте гипотезу $H_0 : \mu_x = \mu_y$ при всех возможных видах конкурирующих гипотез, используя выборку на листе `Стьюдент4` из файла `данные.xlsx`. Положить, что среднеквадратические отклонения не известны, но равны. Напишите скрипт, который это выполняет.

Задание 3. Проверьте гипотезу $H_0 : \mu_x = \mu_y$ при всех возможных видах конкурирующих гипотез, используя выборку на листе `Стьюдент3` из файла `данные.xlsx`. Положить, что среднеквадратические отклонения не известны и не равны. Напишите скрипт, который это выполняет.