**Example for data science.**

Зависимость между случайными величинами – факт интересный и с теоретической, и с практической точек зрения.

Как правило, во всех учебниках приводится пример линейной зависимости, наличие которой обнаруживается с помощью вычисления линейной корреляции, а затем, скажем, методом наименьших квадратов находится и её явный вид.

Но ведь зависимость может быть достаточно простой, но нелинейной.

Поиск такой зависимости более сложная задача, не для нескольких страничек.

Но, чтобы привлечь внимание к этой теме, привожу простой пример, причём в сочетании теории и практики.

Из базы с курсами валют извлекаю курсы американского доллара и евро.

The dependence between random variables is an interesting fact from both theoretical and practical points of view.

As a rule, all textbooks give an example of a linear relationship, the presence of which is detected by calculating a linear correlation, and then, say, its explicit form is found using the least squares method.

But the dependence can be quite simple, but non-linear.

Finding such a dependency is a more difficult task, not for several pages.

But to draw attention to this topic, I give a simple example, moreover, in a combination of theory and practice.

Изображение выглядит как текст

Автоматически созданное описаниеFrom the base with exchange rates I extract the rates of the US dollar and the euro.

Чтобы не загромождать пример большими наборами данных, курсы беру раз в 10 дней.

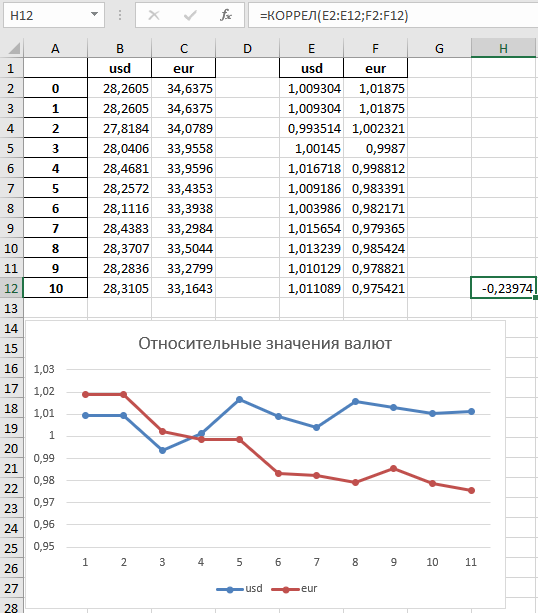
Эта программа создаёт файл curr.xslx с курсами, которые нормирую делением на целую часть курса, а затем изображаю данные в виде графиков. Кроме того, нахожу линейную корреляцию этих данных, которые далее называю относительными значениями валют.

Имею

In order not to clutter up the example with large data sets, I take courses every 10 days.

This program creates a curr.xslx file with courses, which I normalize by dividing by the integer part of the course, and then plot the data. In addition, I find a linear correlation of these data, which I further call the relative values ​​of currencies.

I have

Изображение выглядит как текст

Автоматически созданное описание

Никакой корреляции нет, значит, и линейной зависимости – тоже.

Но замечаю, что возрастание одной валюты, как правило, но не строго, соответствует возрастанию и другой.

Поэтому рассматриваю массив разностей относительных значений валют и нахожу корреляцию уже новых данных.

There is no correlation, which means that there is no linear dependence either.

But I notice that the increase in one currency, as a rule, but not strictly, corresponds to the increase in the other.

Therefore, I consider an array of differences in the relative values ​​of currencies and find a correlation of already new data.

Здесь некоторая корреляция уже имеется.

Изображение выглядит как стол

Автоматически созданное описаниеThere is already some correlation here.

Поскольку пример модельный, я не пытаюсь улучшить зависимость, а нахожу её.

Изображение выглядит как стол

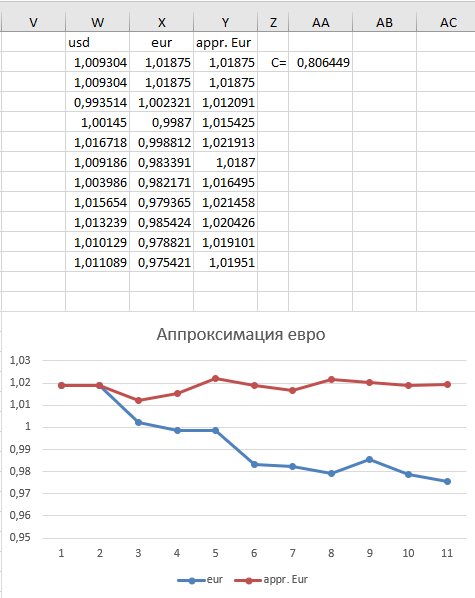
Автоматически созданное описаниеSince the example is model, I am not trying to improve the dependency, but finding it.

Аппроксимация вполне неплохая, а это значит, что если функции доллара и евро обозначить , соответственно, то

The approximation is quite good, which means that if the dollar and euro functions are denoted by respectively, then

где = 0,00945.

Сравниваю настоящее значение евро и полученное из найденной зависимости.

I compare the real value of the euro and obtained from the found dependence.

*Максимальная ошибка аппроксимации – 0,04 от стоимости валюты. Это вполне неплохо.*

***Результат не имеет статистической достоверности, так как на 10 значениях статистика не строится. Напоминаю, это только модельный пример.***

Все вычисления, сделанные в Excel, разумеется, можно реализовать в Pyton.

The maximum approximation error is 0.04 of the currency value. This is quite good.

The result has no statistical significance since statistics are not built on 10 values. Remember, this is just a model example.

All calculations made in Excel, of course, can be implemented in Python.