**Вероятности и статистика (практикум) Домашно 1**

Нека X и Y са последните две цифри на факултетния ви номер, т.е. ???XY

**Зад 1:** Определете променливите съдържащи се в данните "homedata" от пакета "UsingR".

Разгледайте хиляда последователни наблюдения започващи от Y\*500.

За тези данни определете:

а) Най-евтината и най-скъпата къща съответно за 1970 и 2000 година.

б) 5те най-скъпи къщи през 1970г и цената, която имат те през 2000г.

в) Колко са къщите, чиято цена е пораснала с по-малко от 40 000 долара за този период?

г) Представете данните графично, какъв извод може да направите?

д) Съществува ли линейна зависимост между цената в 2000г и тази в 1970?

е) Ако къща е струвала 80 000 през 1970г, каква цена можем да очакваме в 2000г?

ж) Разделете къщите в три категории според цената им през 1970г - евтини до 50 000, нормални с цена от 50 000 до 100 000 и скъпи. Постройте boxplot за цените в 2000г в зависимост от категориите.

з) Пресметнете средната цена на къщите от всяка от трите категории. В коя категория процентно цената е увеличена най-много?

**Зад 2:** Разгледайте данните "Cars93" от пакета "UsingR"

а) преизчислете разходът на автомобилите от мили за галон в литри за сто километра.

б) Как се влияе разходът от броя на цилиндрите? Представете резултата графично.

в) Потърсете връзка между обемът на двигателите и мощността. Определете "outliers", кои модели са това?

Пресметнете корелацията със и без outliers. Направете извод.

г) Представете графично цената в зависимост от броя на пътниците. Направете извод.

**Зад 3.** Генерирайте 50 случайни наблюдения с равномерно разпределение в интервала [0; X+1] и 50 случайни наблюдения експоненциално разпределени с параметър 1/(Y+1)

a) Постройте боксплот за двете извадки на една графика, направете извод.

б) Постройте хистограма за експоненциалните данни. На същата графика начертайте плътността построена по данните и теоретичната плъност.

\*\*\*

Домашното се предава в електронен вариант или на хартиен носител на асистента, при когото сте си избрали да ходите. Решението трябва да съдържа кода на R, резултата от изпълнението му, включително графиките, както и тълкуване, където е необходимо.

\*\*\*