

## Теоретические задачи

Вариант человека определяется по первой букве фамилии (А-З - 0, И-О - 1, П-Я - 2), вариант команды равен сумме вариантов участвующих игроков по модулю 3 + 1.

В каждом варианте требуется решить определенный список подпунктов из задач ниже, однако приветствуется реализация всех подпунктов:

**Вариант 1.** Задача 1 - iii) - а, г, д, е. Задача 2. в), г). Задача 3. а, е.

**Вариант 2.** Задача 1 - i) - в, г, д, е. Задача 2. а), б). Задача 3. б, г.

**Вариант 3.** Задача 1 - ii) - б, в, д, е. Задача 2. б), г). Задача 3. в, д.

1. Проведите тестирование методов хи-квадрат, Колмогорова, Крамера-Мизеса, Андерсона-Дарлинга проверки простой гипотезы для проверки

i) нормальности  $\mathcal{N}(0, 1)$

ii) равномерности  $R[-1.7, 1.7]$

iii) **Стюдентовости с 5 степенями свободы**

для выборок размеров 25, 50, 100:

**а)  $\mathcal{N}(0.1, 1)$**

б)  $R[-1.7, 1.7]$

в) из распределения Лапласа

**г)  $\mathcal{N}(0, 1)$ ,**

**д)  $\mathcal{N}(0, 1)$  при значениях больших 1 по модулю и  $R[-1, 1]$  иначе**

**е) с распределением Стюдента с 5 степенями свободы.**

Для тестирования сгенерируйте по 100 выборок для каждого распределения и подсчитайте для них фактические уровни значимости p-value. Постройте график для э.ф.р. p-value для каждого метода. Для правильного распределения эта э.ф.р. должна быть близка к равномерной ф.р., а для неправильного по возможности быть сильно выше. Какие критерии лучше справились?

2. Смоделировать выборки из  $\mathcal{N}(\mu, \sigma^2)$  и проверить их на нормальность с помощью

а) критерия Колмогорова-Смирнова с оцененными параметрами

б) критерия Лиллефорса

**в) критерия Андерсона-Дарлинга с оцененными параметрами**

**г) критерия Андерсона-Дарлинга для проверки нормальности.**

Моделировать 1000 выборок и построить распределение фактического уровня значимости.

3. Проверить нормальность выборки размеров 20, 50, 100 из

**а)  $\mathcal{N}(0, 1)$**

б)  $\exp(1)$

в) Лапласа с параметром 1

г) хи-квадрат 1

д) стьюдентовские с 4 степенями свободы

**е) фишеровские 3,4**

с помощью критериев Харке-Бера, Шапиро, Д'Агостино.

## Прикладные задачи

### Вариант 1

1. Закон Бенфорда в его житейской интерпретации утверждает, что если выбрать набор каких-то чисел из реальных данных, то распределение первых цифр этих чисел будет иметь вид  $\log_{10}(1+1/k)$ ,  $k = 1, \dots, 9$ . Проверить его **а) на списке стран по населению, б) на числах Фибоначчи в) интересном вам наборе данных.**
2. В файле Rainfall.txt приведены данные о рекордах выпадения дождя в австралийском местечке под названием Туррамура. Можно ли утверждать, что они подчиняются **а) экспоненциальному б) логнормальному распределению?** В обоих случаях использовать и визуальный метод проверки. Если да, то оцените параметры распределения с помощью ОМП.

## Вариант 2

1. Закон Бенфорда в его житейской интерпретации утверждает, что если выбрать набор каких-то чисел из реальных данных, то распределение первых цифр этих чисел будет иметь вид  $\log_{10}(1+1/k)$ ,  $k = 1, \dots, 9$ . Проверить его а) на списке стран по населению, б) на числах Фибоначчи в) интересном вам наборе данных.
2. Kiama Blowhole - это австралийская достопримечательность, представляющая собой отверстие в скале, через которое через некоторые промежутки вылетают фонтаны воды (в русском языке такие объекты называют "дыхалом"). В файле kiama.txt приведены промежутки (в секундах) между такими извержениями, зафиксированные за некоторый промежуток изучения. Можем ли мы утверждать, что извержения образуют пуассоновский процесс (то есть происходят через независимые экспоненциальные времена с одинаковым параметром). Если да, то оцените его параметр с помощью ОМП.

## Вариант 3

1. Закон Бенфорда в его житейской интерпретации утверждает, что если выбрать набор каких-то чисел из реальных данных, то распределение первых цифр этих чисел будет иметь вид  $\log_{10}(1+1/k)$ ,  $k = 1, \dots, 9$ . Проверить его а) на списке стран по населению, б) на числах Фибоначчи в) интересном вам наборе данных.
2. В файле nerve.txt приведены данные о 800 промежутках между прохождением нервных сигналов. Можем ли мы утверждать, что они экспоненциальны с одним и тем же параметром? Если да, то оценить параметр с помощью ОМП.