**Лабораторная работа 2**

Предварительный статистический анализ данных Фишера

(срок сдачи 06.11.2018)

1. Предварительный анализ

Для данных Фишера (iris):

1) проверить выборки, соответствующие переменным Sepal.Length, Sepal.Width, Petal.Length, Petal.Width на независимость и однородность;

2) исследовать закон распределения случайных выборок, соответствующих переменным Sepal.Length, Sepal.Width, Petal.Length, Petal.Width;

3) проверить предположение о нормальном распределении выборок;

4) с помощью графического анализа и тестовых статистик проверить предположение о присутствии в данных аномальных наблюдений;

5) исследовать однородность данных по среднему значению и дисперсии значений переменных Sepal.Length, Sepal.Width, Petal.Length, Petal.Width, относящихся к различным типам ирисов (setosa, versicolor, virginica; переменная Species).

Исследование проводить по индивидуальным наборам данных для указанных переменных и выборок (см. табл.)

**Варианты индивидуальных наборов данных**

|  |  |  |  |
| --- | --- | --- | --- |
| № Варианта | Задание 1.1 | Задание 1.2 | Задание 1.3 |
|  | setosa | setosa | setosa |
|  | setosa | versicolor | versicolor + virginica |
|  | setosa | virginica | versicolor |
|  | versicolor | versicolor | versicolor |
|  | versicolor | virginica | setosa |
|  | versicolor | versicolor + virginica | virginica |
|  | virginica | virginica | virginica |
|  | virginica | setosa | versicolor + virginica |
|  | virginica | versicolor + virginica | versicolor |
|  | versicolor + virginica | versicolor + virginica | versicolor + virginica |
|  | versicolor + virginica | setosa | virginica |
|  | versicolor + virginica | versicolor | setosa |
|  | versicolor | versicolor | versicolor + virginica |

2. Корреляционный анализ

1) осуществить анализ диаграмм рассеяния для каждой пары переменных. Проверить, какие переменные являются положительно коррелированными, отрицательно коррелированными или некоррелированными;

2) построить оценки парных коэффициентов корреляции и проверить их статистическую значимость на уровне ε = 0.05. Сравнить полученные выводы с выводами из пункта 1.

3. Исследование свойств многомерного нормального распределения

1) для набора из 3 переменных, для которых не отклоняется гипотеза о нормальном распределении:

а) проверить предположение, что анализируемые данные являются выборкой из многомерного нормального распределения;

б) в случае выполнения предположения из а) оценить параметры распределения: вектор средних и ковариационную матрицу;

2) по выборке найти оценки условного математического ожидания и условной ковариационной матрицы для случаев (*N* = 3):

а) *х*(1) (зависимая переменная) − первая компонента вектора, а вектор *х*(2) образован из оставшихся двух компонент исходного вектора;

б) *х*(1) образован из первых двух компонент, *х*(2) − это третья компонента исходного вектора.