

Ирисы Фишера

Файзуллина Динара

СПбГУ, 342 группа

Ирисы Фишера — это набор данных для задачи классификации, на примере которого Рональд Фишер в 1936 году продемонстрировал работу разработанного им метода дискриминантного анализа.

Иногда его также называют ирисами Андерсона, так как данные были собраны американским ботаником Эдгаром Андерсоном.

Этот набор данных стал уже классическим, и часто используется в литературе для иллюстрации работы различных статистических алгоритмов.

Ирисы Фишера состоят из данных о 150 экземплярах ириса, по 50 экземпляров из трёх видов — Ирис щетинистый (*Iris setosa*), Ирис виргинский (*Iris virginica*) и Ирис разноцветный (*Iris versicolor*).

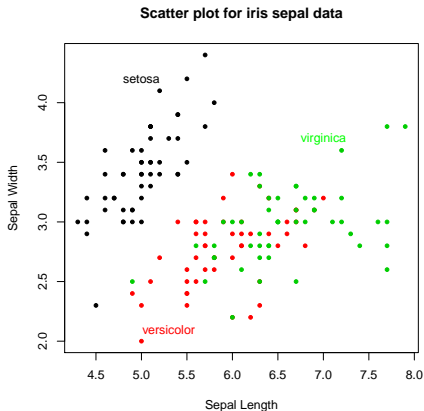
Для каждого экземпляра измерялись четыре характеристики (в сантиметрах):

- Длина наружной доли околоцветника (англ. sepal length)
- Ширина наружной доли околоцветника (англ. sepal width)
- Длина внутренней доли околоцветника (англ. petal length)
- Ширина внутренней доли околоцветника (англ. petal width)

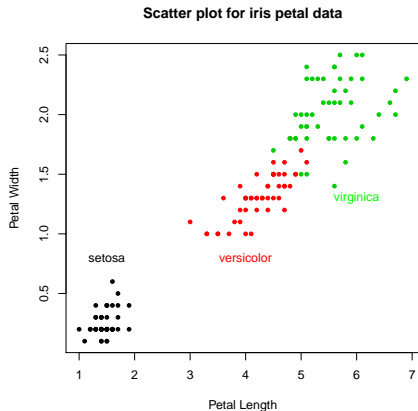
Набор данных

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.10	3.50	1.40	0.20	setosa
2	4.90	3.00	1.40	0.20	setosa
3	4.70	3.20	1.30	0.20	setosa
4	4.60	3.10	1.50	0.20	setosa
5	5.00	3.60	1.40	0.20	setosa
6	5.40	3.90	1.70	0.40	setosa
7	4.60	3.40	1.40	0.30	setosa
8	5.00	3.40	1.50	0.20	setosa
9	4.40	2.90	1.40	0.20	setosa
10	4.90	3.10	1.50	0.10	setosa
11	5.40	3.70	1.50	0.20	setosa
12	4.80	3.40	1.60	0.20	setosa
13	4.80	3.00	1.40	0.10	setosa
14	4.30	3.00	1.10	0.10	setosa
15	5.80	4.00	1.20	0.20	setosa
16	5.70	4.40	1.50	0.40	setosa

Диаграмма рассеивания

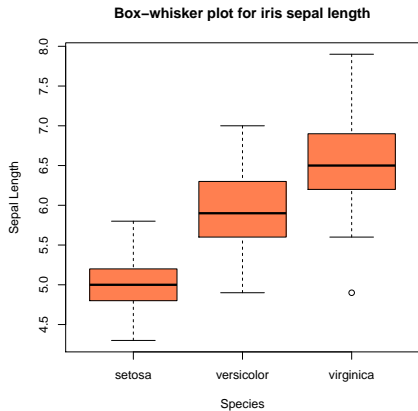


(a) Наружная доля околоцветника

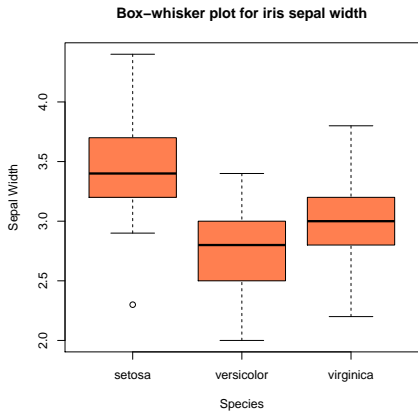


(b) Внутренняя доля околоцветника

Рис.: Диаграммы рассеивания для двух типов характеристик

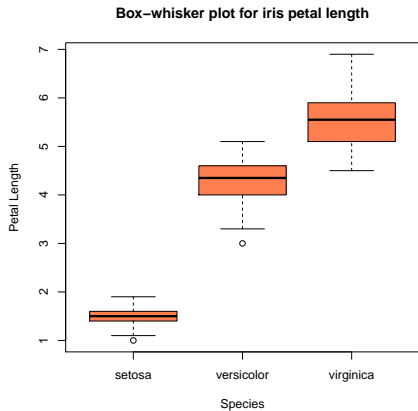


(a) Длина доли



(b) Ширина доли

Рис.: Диаграмма "ящик с усами" для наружной доли околоцветника



(а) Длина доли



(б) Ширина доли

Рис.: Диаграмма "ящик с усами" для внутренней доли околоцветника