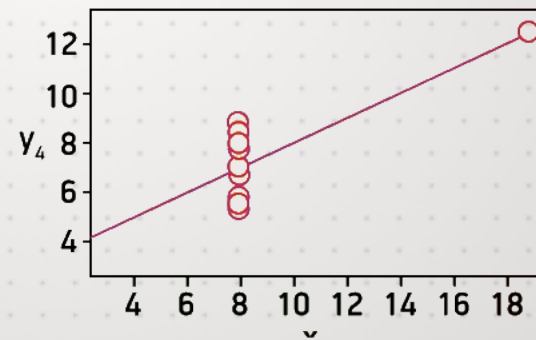
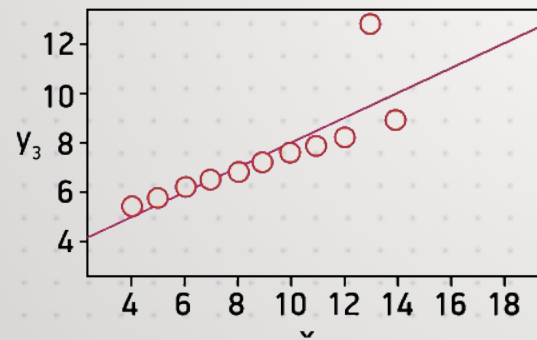
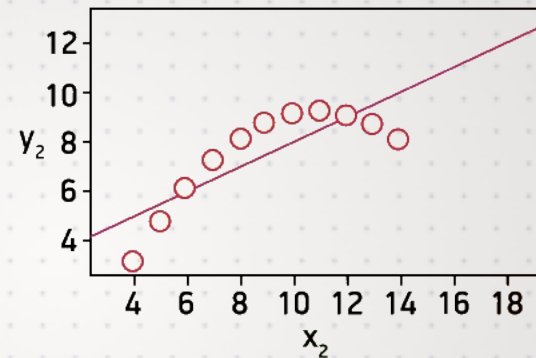
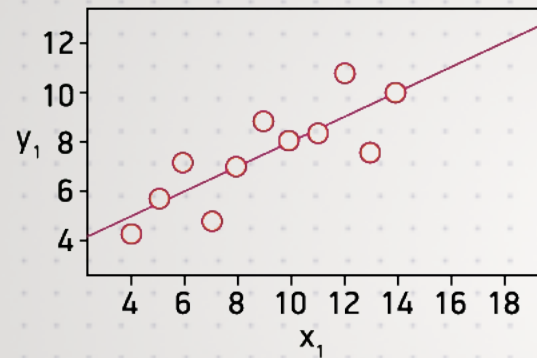


Почему важно визуализировать данные. Квартет Энскомба. Эмпирическая функция распределения

Графический анализ данных

Квартет Энскомба

Квартет Энскомба — четыре набора числовых данных, у которых простые статистические свойства очень похожи, но их графики существенно отличаются.



Значения выборочного среднего и выборочной дисперсии для данных из квартета Энскомба

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = 9.0$$

$$S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = 10.0$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i = 7.5$$

$$S_Y^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2 = 3.75$$

выборочное среднее
выборочная дисперсия

1		2		3		4	
X_1	Y_1	X_2	Y_2	X_3	Y_3	X_4	Y_4
10	8,04	10	9,14	10	7,46	8	6,58
8	6,95	8	8,14	8	6,77	8	5,76
13	7,58	13	8,74	13	12,74	8	7,71
9	8,81	9	8,77	9	7,11	8	8,84
11	8,33	11	9,26	11	7,81	8	8,47
14	9,96	14	8,1	14	8,84	8	7,04
6	7,24	6	6,13	6	6,08	8	5,25
4	4,26	4	3,1	4	5,39	19	12,5
12	10,84	12	9,13	12	8,15	8	5,56
7	4,82	7	7,26	7	6,42	8	7,91
5	5,68	5	4,74	5	5,73	8	6,89
9	7,5	9	7,5	9	7,5	9	7,5
10	3,75	10	3,75	10	3,75	10	3,75

Эмпирическая функция распределения

Эмпирическая функция распределения — выборочный аналог функции распределения, который мы можем построить по имеющейся у нас выборке.

Функция распределения показывает вероятность попадания случайной величины в интервал от $-\infty$ до x .

$$F_{\xi}(x) = P\{\xi \leq x\}, x \in \mathbb{R}$$

Эмпирическая функция распределения

$$F_n(x) = \frac{\mu_n(x)}{n}$$

$\mu_n(x)$ — количество наблюдений $\leq x$