

ДЗ №1 по регрессии.

1. Построить генератор данных по зашумленной зависимости $y = \beta x + \varepsilon$ по коэффициентам β и дисперсии шума своего варианта. Задать 3 уровня регрессора x : 200, 220, 240. Объем выборки (количество пар x, y) – 10.
2. С помощью функции regress построить оценки коэффициентов регрессии $\hat{\beta}$.
3. Построить истинную линию зависимости $y(x)$ с исходными коэффициентами β . На том же графике построить восстановленную линию $y(x)$ по оценкам коэффициентов регрессии $\hat{\beta}$. На график также вывести точки сгенерированных исходных данных.

ДЗ №2 по регрессии

1. Методом статистического моделирования построить распределения оценок коэффициентов регрессии как случайных величин. По результатам численного моделирования определить 95% доверительный интервал коэффициентов.
2. Методом статистического моделирования построить полигон относительных частот прогноза $\hat{y}(x_n)$ при $x=200$ и $x = 240$.
3. Методом статистического моделирования построить полигон относительных частот ошибки предсказания истинного значения $e_n = \hat{y}(x) - y(x)$ при $x=200$ и $x = 300$.
4. Методом статистического моделирования построить полигон относительных частот распределения ошибки предсказания нового измерения $e_n = \hat{y}(x) - y_n$ при $x=200$ и $x = 300$.
5. Построить теоретические плотности распределения ошибки предсказания нового измерения e_n при $x=200$ и $x = 300$. Отметить 95%-доверительный диапазон. График совместить с графиками из предыдущего пункта.

ДЗ №3 по регрессии

1. Задаться объемом выборки n из диапазона $n > 10$.
2. Смоделировать N выборок объема n , причем для каждой выборки сгенерировать одно *новое* значение, задавшись $x = 220$.
3. По каждой выборке восстановить регрессионную модель.
4. Для каждой из N построенных регрессионных моделей, используя распределение Стьюдента рассчитать 95%-доверительный интервал прогноза нового значения.
5. Рассчитать ширину доверительного интервала (разница между верхней и нижней границей интервалов).
6. Для всех N выборок рассчитать долю выборок, где новое значение вышло за доверительный интервал прогноза.
7. Увеличить объем выборки n в два раза. Убедиться, что средняя ширина доверительного интервал сузилась, но доля выходов за доверительный интервал не изменилась.