

Learning outcomes к проверочной работе 5 марта

1. Множественная линейная регрессия. Ключевые предикторы и контрольные переменные. Каким требованиям должны удовлетворять контрольные переменные?
2. Интерпретация коэффициентов в множественной линейной регрессии
3. ANOVA таблица для множественной линейной регрессии: уметь заполнить пропуски и сделать выводы на основе статистик
4. Запись спецификации регрессионной модели в векторно-матричном виде
5. Выведение формулы оценки коэффициентов: $(X^T X)^{-1} X^T y$ (релевантно как для случая парной, так и множественной регрессии)
6. Уметь по заданным значениям предикторов и значениям отклика получить вектор оценок коэффициентов регрессионной модели

Пример 1. Ниже в таблице представлены значения переменных: X , Z , Y .

X	-2	2	1	-1	0
Z	2	0	0	-1	-1
Y	2	6	3	2	10

Без использования R получите оценки коэффициентов:

- (a) в регрессии Y на X
- (b) в регрессии Y на X и Z (то есть, X и Z одновременно выступают предикторами – случай множественной регрессии)

Представьте промежуточные расчеты, выпишите полученный вектор оценок коэффициентов и запишите спецификацию модели, подставив эти оценки в уравнение.

Примечание: в проверочной работе будет только случай с одним предиктором, как в первом пункте и на семинаре

7. Мультиколлинеарность: определение, примеры
8. Откуда берется мультиколлинеарность?
9. Последствия мультиколлинеарности
10. Уметь объяснить и показать, почему в случае строгой мультиколлинеарности оценки в модели не могут быть получены
11. Почему в случае сильной мультиколлинеарности оценки модели являются неустойчивыми?
12. Способы диагностики мультиколлинеарности:
 - (a) исходная корреляционная матрица предикторов
 - (b) визуализация: scatterplots
 - (c) дополнительно можно посмотреть на частные корреляции
 - (d) Variance inflation factor больше 10 указывает на сильную мультиколлинеарность (либо tolerance: меньше 0.1 указывает на сильную мультиколлинеарность)
13. Гетероскедастичность: определение, примеры
14. Откуда берется гетероскедастичность?

15. Последствия гетероскедастичности

16. Способы выявления гетероскедастичности:

- (a) теоретические предпосылки
- (b) визуализация
- (c) формальные тесты: тест Уайта и тест Goldfeld-Quandt