

## Learning outcomes №2

1.  $R^2$  скорректированный: формула для расчета, отличие от классического коэффициента детерминации
2. Кросс-валидация в контексте регрессионного анализа
3. Сравнение вложенных линейных регрессионных моделей при помощи F-теста
4. Информационные критерии AIC, BIC: знать формулы, как рассчитать информационные критерии, уметь делать вывод, знать, какой штраф накладывается за добавление новых параметров
5. Линейная вероятностная модель и ее ограничения
6. Представление модели бинарного выбора:
  - подход, основанный на представлении зависимой переменной как латентной
  - подход, использующий переход от вероятности к шансам
7. Вероятность того, что  $Y$  принимает значение 1, как функция распределения от предсказанной части модели
8. Допущения об ошибках: логит- и пробит-модель. Знать функцию стандартного логистического распределения и уметь ее применять
9. Сигмоида: понимать, почему такой характер зависимости вероятности того, что  $Y$  принимает значение 1, от значения предиктора
10. Интерпретация оценок коэффициентов:
  - исходные (в терминах склонности)
  - через отношения шансов
  - через предельные эффекты
11. Уметь переводить исходные оценки модели в отношения шансов
12. Тест Hosmer-Lemeshow: какую гипотезу проверяем, логика реализации теста, критика
13. Понимать, как устроена confusion matrix
14. Уметь считать по confusion matrix (и, разумеется, понимать, что это за величины, что они показывают)
  - ошибку первого рода
  - ошибку второго рода
  - мощность критерия
  - чувствительность
  - специфичность
  - ассигасу и сравнивать с baseline ассигасу
15. Понимать дилемму соотношения специфичности и чувствительности
16. ROC-curve: как устроен график, что показывает AUC (площадь под кривой)
17. Сравнение вложенных моделей посредством likelihood-ratio test (тест отношения правдоподобия)

18. Представление спецификации логистической модели с порядковым откликом
  - подход, основанный на представлении зависимой переменной как латентной
  - подход, использующий переход от вероятности к шансам
19. Уметь рассчитывать вероятность того, что  $Y$  принимает конкретное значение
20. Уметь рассчитывать функцию распределения от конкретного значения категории
21. Допущение о параллельности регрессий: смысл допущения, последствия нарушения допущения, тестирование допущения
22. Multiclass ROC-curve и AUC: в чем ограничение усреднения AUC, полученных по результатам оценивания промежуточных моделей с бинарным откликом (сравнение категорий попарно)?
23. Представление спецификации мультиномиальной модели через набор моделей с бинарным откликом
24. LSDV (least-squared-dummy-variables) модель: представление фиксированных эффектов через дамми-переменные. Классическая спецификация с фиксированными эффектами для моделирования различий в «стартовых» условиях. Интерпретация оценок коэффициентов в этой модели
25. Можно ли полагаться на  $R^2$  как меру качества в LSDV-модели?
26. Модель с внутригрупповым преобразованием. Каким образом осуществляется внутригрупповое преобразование и в каких случаях оно удобно?
27. Что «содержательно» включают в себя дамми-переменные (какие характеристики)? Можно ли говорить, что модель с фиксированными эффектами полностью исключает проблему эндогенности? Объясните свой ответ на примерах
28. «Техника» выведения оценок коэффициента при предикторе в модели с фиксированными эффектами. Оценка коэффициента как сумма взвешенных коэффициентов из моделей, оцененных отдельно по подвыборкам
29. F-test для выбора между объединенной (pooled) моделью регрессии (без учета неоднородности) и моделью с фиксированными эффектами. Интерпретация результатов теста
30. Почему невозможность оценить коэффициенты при неизменяющихся во времени предикторах в модели с фиксированными эффектами при анализе панельных данных не является проблемой? Почему нам в этом случае не нужны неизменяющиеся во времени предикторы?
31. Модель со случайными эффектами: спецификация, допущения, ограничения, метод оценивания GLS (generalized least squares), FGLS (feasible generalized least squares)
32. Тест Хаусмана: нулевая гипотеза и альтернатива, распределение статистики, интерпретация результатов
33. Модель с фиксированными эффектами на временные периоды. Интерпретация
34. Two-ways model (с включением эффектов как на пространственные, так и временные периоды): что отражают оценки коэффициентов, в чем сложность интерпретации
35. Модель с включением переменных взаимодействия между дамми-переменными и ключевым предиктором для оценивания различий в характере взаимосвязи этого ключевого предиктора и зависимой переменной. Интерпретация оценок коэффициентов в этой модели.
36. Модель разность разностей (DiD):
  - уметь считать оценки коэффициентов для модели DiD (простейшая спецификация без контрольных переменных) на основании таблицы со средними значениями для группы воздействия и контрольной группы в периоды ДО и ПОСЛЕ введения воздействия

- уметь интерпретировать соответствующие оценки коэффициентов
- допущение параллельности трендов
- понятие counterfactual outcome (гипотетический исход) в рамках модели DiD