

Семинар 7.

1. Набор данных *crime.dta* содержит информация об арестах в течение 1986 года по 2725 мужчинам, родившимся в Калифорнии в 1960 или 1961 гг. Каждый из мужчин в выборке был арестован по крайней мере однажды до 1986 года. Вопрос: что объясняет, какое количество раз мужчина совершает преступления? Имеются следующие данные:

- *par86* – количество арестов в 1986 г.
- *rcnv* – доля предыдущих арестов, приведших к осуждению (прокси для неотвратимости наказания)
- *avgsen* – среднее время заключения по предыдущим случаям заключения (в месяцах) (прокси для суровости наказания)
- *tottime* – общее время, проведенное в тюрьме после достижения возраста 18 лет (в месяцах)
- *ptime86* – число месяцев в тюрьме во время 1986 г. (не может быть арестован, пока в тюрьме)
- *qemp86* – количество кварталов, в которых имел работу в течение 1986 г. (возможности на рынке труда)
- *inc86* – легальный доход, 1986, 100\$
- *durat* – длительность последнего периода безработицы (в месяцах)
- *black* =1, если афроамериканец
- *hispan* =1, если латиноамериканец
- *born60* =1, если родился в 1960 г.
- *crime86* =1, если был арестован хотя бы однажды в 1986 г.

2. Набор данных *cornwell&rupert.csv* содержит информацию о 595 индивидуумах за 7 лет по следующим показателям:

- *exp* – опыт работы
- *wks* – отработанные недели
- *occ* = 1, если "синие воротнички"
- *ind* = 1, если обрабатывающая промышленность
- *south* =1, если проживает на юге
- *smsa* =1, если проживает в городе (SMSA)
- *mc* =1, если женат
- *fem* =1, если женщина

- $union = 1$, если заработная плата установлена профсоюзным договором
- ed – годы образования
- $lwage$ – логарифм заработной платы

Оцените следующие модели регрессии для логарифма заработной платы: pooled, FE и RE. С помощью соответствующих тестов выберите лучшую из оцененных моделей.