Семинар 8.

1. В файле *Chow.xls* содержатся данные об экономике Баккардии в период с 1 квартала 2015 года по 4 квартал 2022 года. Показатели выражены в миллиардах баккардийских крон 2015 года.

   a) Оцените следующую модель регрессии:

   $$C_t = \beta_1 + \beta_2 Y_t + \beta_3 D2_t + \beta_4 D3_t + \beta_5 D4_t + \varepsilon_t,$$

   где C_t – конечное потребление в момент времени t, Y_t – конечное потребление в момент времени t, Y_t – располагаемый доход в момент времени t, Dj_t – дамми переменная на квартал ($j = 2, 3, 4$).
   Проинтерпретируйте полученные результаты.

   b) На уровне значимости 5% проверьте гипотезу о наличии сезонности. Сформулируйте нулевую и альтернативную гипотезы.

   c) Оцените модель в следующем виде:

   $$C_t = \beta_1 Y_t + \beta_2 D1_t + \beta_3 D2_t + \beta_4 D3_t + \beta_5 D4_t + \varepsilon_t.$$

   Сравните полученные результаты с предыдущим пунктом.

   d) Попробуйте улучшить модель, включив в нее переменные взаимодействия:

   $$C_t = \beta_1 + \beta_2 Y_t + \beta_3 D2_t + \beta_4 D3_t + \beta_5 D4_t + \beta_6 (Y * D2) + \beta_7 (Y * D3) + \beta_8 (Y * D4) + \varepsilon_t.$$

   Проинтерпретируйте полученные результаты.

2. В файле *Chow.xls* содержатся данные об экономике Баккардии в период с 1 квартала 2015 года по 4 квартал 2022 года. Используя известные Вам показатели, проанализируйте наличие в данных влиятельных наблюдений и выбросов.

3. Всего имеется 100 наблюдений. Для первых 50-ти наблюдений

   $$X'X = \begin{pmatrix} 50 & 300 \\ 300 & 2100 \end{pmatrix},\ X'y = \begin{pmatrix} 300 & 2000 \end{pmatrix}',\ y'y = 2100.$$

   По последним 50-ти наблюдениям:

   $$X'X = \begin{pmatrix} 50 & 300 \\ 300 & 2100 \end{pmatrix},\ X'y = \begin{pmatrix} 300 & 2200 \end{pmatrix}',\ y'y = 2500.$$

По первым 50-ти наблюдениям оценивается модель $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, по последним 50-ти наблюдениям оценивается модель $y_i = \gamma_1 + \gamma_2 x_i + \varepsilon_i$. Предположеним, что во всех 100 наблюдениях ε_i независимы и нормальны $N(0;\sigma^2)$. На уровне значимости 5% проверьте гипотезу $H_0: \beta = \gamma$.

Посчитаем RSS в каждой из моделей

$$RSS = y'(I - X(X'X)^{-1}X')y = y'y - y'X(X'X)^{-1}X'y = y'y - (X'y)'(X'X)^{-1}X'y$$

$RSS_1 = 2000 - 1933.33 = \frac{500}{3}$

$RSS_2 = 2500 - 2333.33 = \frac{500}{3}$

$X'X = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}$, $X'y = \begin{pmatrix} \sum_{i=1}^{100} y_i \\ \sum_{i=1}^{100} x_i y_i \end{pmatrix}$ значит, новый $X'X = \begin{pmatrix} 100 & 600 \\ 600 & 4200 \end{pmatrix}$ а новый $X'y = \begin{pmatrix} 600 \\ 4200 \end{pmatrix}$.

$y'y = \sum_{i=1}^{100} y_i^2 = (y'y)_{1st} + (y'y)_{2nd} = 4600$

$RSS_{pooled} = 4600 - 4200 = 400$

Тест Чоу $\frac{(RSS_{pooled} - RSS_1 - RSS_2)/k}{(RSS_1 + RSS_2)/(n-2k)} = \frac{(400-500/3-500/3)/2}{1000/3/96} = \frac{96}{10} = 9.6 > 3.09 = F_{2,96}$ гипотеза о том, что $\beta = \gamma$ отвергается.