

## Тема 7: "Прогнозирование значений эндогенной переменной, проверка её адекватности и ошибки спецификации"

### Задание 1.

- Используя 90% выборки постройте множественную модель линейной регрессии:  

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$$
- Проверьте ее адекватность на основе оставшихся 10% наблюдений.
- Изучите выборку на однородность, используя тест Чоу.
- Постройте модель парной регрессии с наиболее влияющим фактором. Сравните исходную модель с парной моделью. Выберите лучшую.
- Выбранную модель проверьте на наличие пропущенных переменных с помощью теста Рамсея.
- Выполните задания, используя инструменты R и Python.

Объем денежных накоплений в РФ миллиард руб. (y)	Сводные данные о продаже наличной иностранной валюты кредитными организациями физическим лицам (x1)	Курс рубля к доллару США (x2)	Цена нефти марки "Юралс" (x3)	Динамика потребительских цен по группам товаров и услуг (месяц к соответствующему месяцу предыдущего года,%) (x4)
12 345,00	2769	29,67	94,2	9,6
12 091,90	3065	28,94	101,1	9,5
12 339,70	3899	28,43	111,6	9,5
12 450,60	3790	27,5	119,7	9,6
12 832,10	3963	28,07	112	9,6
12 892,10	4224	28,08	112	9,4
13 224,50	4645	27,68	115,3	9
13 449,90	4914	28,86	109,6	8,2
13 476,80	4830	31,88	112,6	7,2
13 588,90	4815	29,9	108,7	7,2
13 600,40	4401	31,32	110,9	6,8
13 768,50	5189	32,2	108	6,1
14 796,60	3434	30,36	109,8	4,2
14 424,20	5195	28,95	119,2	3,7
14 607,30	5401	29,33	123,3	3,7
14 688,00	4239	29,36	117,8	3,6
15 057,30	4558	32,45	109,2	3,6
15 161,00	4675	32,82	93,5	4,3
15 534,60	6558	32,19	102,7	5,6
15 544,50	6253	32,92	113,5	5,9
15 644,40	5410	30,92	112	6,6
15 717,80	5023	31,53	110,8	6,5

## Задание 2.

Решите приведенные ниже задачи.

1. Сравниваются две модели линейной регрессии при помощи F-теста:

модель 1:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \varepsilon_t, \quad RSS_1 = 10, \quad t = 1, \dots, 55,$$

модель 2:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t, \quad RSS_2 = 60, \quad t = 1, \dots, 55.$$

Какая из моделей предпочтительнее, если  $F_{kp} = 3,18$  при  $\alpha = 0,05$ .

Какие аргументы использованы для определения  $F_{kp}$ ?

2. На основе квартальных данных с 1971 по 1976 г. с помощью МНК получено следующее уравнение:

$$\hat{Y}_t = \underset{(2,14)}{1,12} - \underset{(0,0034)}{0,0098} \cdot X_{t1} - \underset{(3,42)}{5,62} \cdot X_{t2} + \underset{(0,009)}{0,044} \cdot X_{t3}, \quad ESS=110,00;$$

$RSS=20,8$ .

Для той же исходной модели были отдельно проведены две регрессии на основе данных:

1) с 1-го квартала 1971 г. по 1-ый квартал 1975 г.,  $RSS_1 = 14,0$ ;

2) со 2-го квартала 1975 г. по 4-ый квартал 1976 г.,  $RSS_2 = 2,0$

При помощи теста Чоу проверьте гипотезу о том, что между 1-м и 2-м кварталами 1975 г. произошло структурное изменение,  $F_{kp} = 3$ , при  $\alpha = 0,05$ .

По каким аргументам определено критическое значение статистики Чоу?

3. Построить доверительный интервал для эндогенной переменной  $Y_6$  по результатам оценивания выборки из 5 наблюдений, проверить адекватность модели:

$$\hat{Y}_t = \underset{(S_{\hat{\beta}_1})}{\hat{\beta}_1} + \underset{(S_{\hat{\beta}_2})}{\hat{\beta}_2} \cdot X_t = \underset{(1,5)}{-2,0} + \underset{(0,05)}{0,14} \cdot X_t, \quad t_{\beta_1} = -1,33, \quad t_{\beta_2} = 2,8, \quad t_{kp} = 3,$$

$$Y_6 = 2,5, \quad X_6 = 30, \quad \sqrt{\text{var}(Y_6 - \hat{Y}_6)} = 1,2.$$