

Семинар 2

Задание 1. По указанным ниже значениям предиктора и отклика (x и y соответственно) найдите вектор оценок коэффициентов в регрессионной модели y на x . Используйте для этого общую формулу оценки коэффициентов в векторно-матричном виде (релевантную как для парной, так и для множественной регрессии). Запишите промежуточные расчеты. В качестве ответа запишите сам вектор и полную спецификацию модели, подставив эти оценки коэффициентов.

x	4	0	2	2	1
y	2	1	1	5	2

Задание 2. Ниже представлены результаты анализа разложения вариации по линейной парной регрессионной модели, построенной по выборке из 15 наблюдений.

Analysis of Variance Table

```
Response: y
      df  sum_sq mean_sq    f    PR(>F)
x      ...    ...     ...  0.6526     ...
Residual ... 56.116     ...
```

1. Восстановите пропуски в таблице
2. Вычислите коэффициент детерминации и проинтерпретируйте его значение
3. На основе указанной выдачи сделайте вывод о (не)значимости коэффициента детерминации, свой ответ поясните

Задание 3.

На данных по 44 городам построена модель, объясняющая динамику уровня преступности за последние 10 лет. `change_in_crime_rate` — прирост преступности в %, `change_in_pop` — прирост численности населения, %; `kids` — процент детей; `free_lunch` — процент бесплатных школьных обедов; `income_change` — прирост доходов домохозяйств.

```
Coefficients:
      coef      std. err   t  Pr>|t|  [0.025  0.975]
Intercept  -22.3548     12.3097
change_in_pop  0.3188     0.2052
kids         1.1128     0.2869
free_lunch   -0.3681     0.0973
income_change -0.1944     0.3681
---
```

```
      df sum_sq mean_sq    f    PR(>F)
change_in_pop    803.2
kids             1380.1
free_lunch       3186.6
income_change     60.6
Residual        8476.0
---
```

- Восстановите все пропуски в таблице
- Определите статистически значимые оценки коэффициентов
- Проинтерпретируйте оценку константы

- Проинтерпретируйте оценку коэффициента при предикторе «процент бесплатных школьных обедов»
- Рассчитайте коэффициент детерминации и проверьте гипотезу о том, что регрессия на константу (то есть, модель без объясняющих переменных) не хуже модели с предикторами. Сформулируйте нулевую и альтернативную гипотезы для критерия, а также содержательный вывод