

Домашнее задание 4:
тренировка перед проверочной работой

Задание 1. Доделайте Задание 1 из Семинарского листочка 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 value Pr(>F)
x      ...    ...      ...    0.6526   ...
Residuals ... 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. Для регрессионной модели, представленной в задании 1, оценка коэффициента при предикторе составляет 0.15, а оценка константы равна 4. Ниже также дана ковариационная матрица коэффициентов для указанной модели.

	(Intercept)	x
(Intercept)	1.6111	-0.2068
x	-0.2068	0.0323

1. В общем виде проинтерпретируйте оценки коэффициентов.
2. На основе значений этой ковариационной матрицы проверьте, значимы ли коэффициенты при предикторе и константа на основании p-value. Для этого выпишите нулевую гипотезу и альтернативу для каждого коэффициента, статистики для проверки гипотез, распределение статистик при верной нулевой гипотезе и значения p-value. Сделайте вывод.
3. Постройте 90%-ые доверительные интервалы для оценок коэффициентов и проинтерпретируйте их (прокомментируйте, что полученные доверительные интервалы говорят о (не)значимости коэффициентов).

Задание 4. Покажите, что $\text{Corr}^2(\hat{y}_i, y_i) = R^2$.

Задание 5. На основе представленных ниже данных оценивается парная линейная регрессия y на x .

y	x
7	2
4	4
5	3
10	6
9	7
12	9
8	3
9	6

1. Вручную рассчитайте оценки коэффициентов константы и коэффициента при предикторе. Запишите спецификацию модели и проинтерпретируйте в общем виде полученные оценки.
2. Чему равен коэффициент корреляции Пирсона между зависимой переменной и предсказанным на основе модели (см. условие выше) значением отклика? Ответьте на поставленный вопрос без предварительных расчетов предсказанных значений зависимой переменной.
Hint: обратите внимание на то, что это парная модель регрессии, поэтому в данном случае необязательно получать при ответе на этот вопрос предсказанные значения, можно ориентироваться на сам предиктор и его корреляцию с зависимой переменной
3. Рассчитайте коэффициент детерминации для указанной регрессионной модели. Проинтерпретируйте полученное значение.
4. Проверьте значимость коэффициента детерминации.
5. Если переоценить модель с использованием вместо исходного предиктора переменную – результат деления x на 10, каким образом изменятся оценки коэффициентов в модели? Придумайте практическую задачу, когда может быть содержательно осмысленно подобное преобразование предиктора.