

1. Храбрый исследователь Вениамин поделил выборку на обучающую  $(X, y)$  и тестовую  $(X_{test}, y_{test})$ . Регрессоры  $X$  и  $X_{test}$  Вениамин считает нестохастическими, а предпосылки теоремы Гаусса-Маркова — выполненными на всей исходной выборке. Естественно,  $\hat{y}_{test} = X_{test}\hat{\beta}$ , где  $\hat{\beta}$  оценивается по обучающей выборке.

Помогите Вениамину найти  $\text{Var}(\hat{y}_{test})$  и  $\text{Cov}(\hat{\beta}, \hat{y}_{test})$ .

2. Рассмотрим матрицу  $X$  полного ранга с  $n$  наблюдениями и  $k$  столбцами. В каких границах могут лежать диагональные элементы матрицы-шляпницы  $H$ ? Чему равно их среднее значение? Подсказка: найдите  $\text{Var}(\hat{y})$  и  $\text{Var}(\hat{u})$  в рамках предпосылок теоремы Гаусса-Маркова.

3. Рассмотрим стандартный  $t$ -тест на равенство некоторого коэффициента бета нулю. Докажите, что

$$t^2 = \frac{RSS_r - RSS_{ur}}{RSS_{ur}/(n - k)},$$

где  $RSS_r$  — сумма квадратов остатков в модели без тестируемого коэффициента (выкинут регрессор при проверяемом коэффициенте),  $RSS_{ur}$  — аналогичная сумма в модели с включённым тестируемым коэффициентом,  $k$  — число оцениваемых коэффициентов бета в модели с тестируемым коэффициентом,  $n$  — количество наблюдений.

Утешительный приз: упростите эту формулу для случая парной регрессии и докажите её :)

4. Рассмотрим стандартную ошибку оценки коэффициента бета при регрессоре  $z$  в множественной регрессии. Докажите, что

$$se^2(\hat{\beta}_z) = \frac{RSS/(n - k)}{\sum (z_i - \bar{z})^2} \cdot \frac{1}{1 - R_z^2},$$

где  $R_z^2$  — коэффициент детерминации во вспомогательной регрессии объясняющей переменной  $z$  на остальные объясняющие переменные.

Утешительный приз: упростите эту формулу для случая парной регрессии и докажите её :)

5. Винни-Пух нашёл случайный вектор  $w$  и одномерную случайную величину  $z$ . Также он узнал, что  $\text{Corr}(w, w) = A$  и  $\text{Corr}(w, z) = b$ .

К сожалению, у Винни-Пуха опилки в голове, а он очень хочет найти такую линейную комбинацию компонент вектора  $w$ , которая была бы сильнее всего коррелирована со случайной величиной  $z$ .

Помогите Винни-Пуху!

Как выглядят веса этой линейной комбинации? Чему равна максимально возможная корреляция?

6. Машенька построила парную регрессию по 11 наблюдениям с  $R^2 = 0.95$ . Чтобы напакостить Машеньке, Вовочка переставил в случайном порядке значения зависимой переменной и предложил Машеньке заново оценить модель.

Какой ожидаемый  $R^2$  получит Машенька?