Задача 1. Рассматривается модель $y_i = \alpha + \beta x_i + \varepsilon_i$, $\mathbb{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$. При каких значениях c_i несмещенная оценка $\hat{\beta} = \frac{\sum_{i=})^n c_i y_i}{\sum_{i=1}^n c_i x_i}$ имеет наименьшую дисперсию?

Ответ: $c_i = x_i - \overline{x}$

Задача 2. Рассматривается модель $y_i = \alpha + \beta x_i + \varepsilon_i$, $\mathbb{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$. При каких значениях c_i несмещенная оценка $\hat{\beta} = \frac{\sum_{i=1}^n c_i (y_i - \overline{y})}{\sum_{i=1}^n c_i (x_i - \bar{i})}$ имеет наименьшую дисперсию?

Ответ: $c_i = x_i - \overline{x}$

Задача 3. Рассматривается модель $y = X\beta + \varepsilon$, где $\mathbb{E}(\varepsilon) = 0$, $\mathrm{Cov}(\varepsilon) = \sigma^2 I$. Найдите значение параметра c, при котором несмещенная оценка $\hat{\beta} = ((X^T X)^{-1} + cI)y$ имеет наименьшую ковариационную матрицу.

Задача 4. Рассматривается модель $y = X\beta + \varepsilon$, где $\mathbb{E}(\varepsilon) = 0$, $\mathrm{Cov}(\varepsilon) = \sigma^2 I$. Найдите такую матрицу C, для которой несмещенная оценка $\hat{\beta} = ((X^T X)^{-1} + C)y$ имеет наименьшую ковариационную матрицу.

Задача 5. Пусть $y_i = \alpha + \varepsilon_i$, $\mathbb{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$ и $\hat{\alpha}$ — МНК-оценка неизвестного параметра α. Найдите $\mathbb{E}(\hat{\alpha})$ и $\mathrm{Var}(\hat{\alpha})$

Задача 6. Пусть $y_i = \beta x_i + \varepsilon_i$, $\mathbb{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$ и $\hat{\beta}$ — МНК-оценка неизвестного параметра β. Найдите $\mathbb{E}(\hat{\beta})$ и $\mathrm{Var}(\hat{\beta})$

Задача 7. Пусть $y_i = \alpha + \beta x_i + \varepsilon_i$, $\mathbb{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ при $i \neq j$, $\hat{\alpha}$ и $\hat{\beta}$ — МНК-оценки неизвестных параметров α и β соответственно. Найдите

(a) $\mathbb{E}(\hat{\alpha})$ и $\mathrm{Var}(\hat{\alpha})$

(b) $\mathbb{E}(\hat{\beta})$ и $\mathrm{Var}(\hat{\beta})$

Задача 8 При помощи метода наименьших квадратов оценивается модель $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ по 23 наблюдениям. Результаты оценивания:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ \widehat{\mathrm{Cov}} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 5 \end{bmatrix}$$

На уровне значимости 5% проверьте гипотезу $H_0 : \beta_2 = 0$ против альтернативы $H_1 : \beta_2 > 0$. Выпишите:

(a) Тестовую статистику (формулу)

(b) Распределение тестовой статистики

(c) Наблюдаемое значение тестовой статистики

(d) Область, в которой H_0 не отвергается

(e) Статистический вывод

На уровне значимости 5% проверьте гипотезу $H_0 : 2\alpha = \beta_1$ против альтернативы $H_1 : 2\alpha \neq \beta_1$. Выпишите:

(f) Тестовую статистику (формулу)

(g) Распределение тестовой статистики

(h) Наблюдаемое значение тестовой статистики

(i) Область, в которой H_0 не отвергается

(j) Статистический вывод

Задача 9. Пусть регрессионная модель $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = \begin{bmatrix} \alpha & \beta_1 & \beta_2 \end{bmatrix}^T$ и $\varepsilon \sim N(0, \sigma^2 I)$. Известно, что

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \; X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \; (X^T X)^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 4/3 & -1/3 \\ 0 & -1/3 & 4/3 \end{bmatrix}$$

(a) Укажите число наблюдений n

(b) Укажите число регрессоров с учётом свободного члена

(c) Найдите TSS

(d) Найдите $\hat{\beta}$

(e) Найдите RSS

(f) Найдите ESS

(g) Найдите R^2

(h) Найдите R^2_{adj}

(i) Найдите $\hat{\sigma}^2$

(j) Постройте 80%-ый доверительный интервал для $\hat{\sigma}^2$

(k) Постройте 80%-ый доверительный интервал для σ

(l) Найдите $\widehat{\mathrm{Cov}}(\hat{\beta})$ — оценку ковариационной матрицы для $\hat{\beta}$

(m) Постройте 90%-ый доверительный интервал для α

(n) Постройте 95%-ый доверительный интервал для $\alpha + \beta_1 + \beta_2$

(o) Постройте 95%-ый доверительный интервал для $\alpha + 2\beta_1 + 3\beta_2$

Задача 10. Рассматривается модель

$$\text{Цена}_i = \alpha + \beta_1 \text{Пробег}_i + \beta_2 \text{ЧП}_i + \beta_3 \text{Иномарка}_i + \beta_4 \text{Литраж}_i + \beta_5 \text{Длина}_i + \varepsilon_i$$

где переменная *ЧП* означает число поломок, а переменная *Иномарка* равна 1, если автомобиль произведён за рубежом и равна 0 в противном случае. Исследователь предполагает, что ожидаемая цена иномарки, потерпевшей четыре поломки, совпадает с ценой ни разу не ломавшегося отечественного автомобиля с аналогичными характеристиками. Сформулируйте эту гипотезу в терминах коэффициентов регрессии.

Задача 11. Изучается зависимость уровня годового дохода (переменная *ГД* в сотнях тыс. руб.) финансового аналитика в зависимости от опыта работы OP (годы), пола (переменная *Пол* равная 1 для мужчин и 0 для женщин), владения английским языком (переменная FE равна 1, если аналитик свободно владеет английским языком и 0 в противном случае) и наличия сертификата CFA (переменная *CFA* равна 1, если сертификат есть). Оцененная модель имеет вид:

$$\widehat{\text{ГД}} = 1200 + 200 \cdot \text{ОП} - 10 \cdot \text{ОП}^2 + 200 \cdot \text{Пол} + 400 \cdot \text{FE} + 150 \cdot \text{CFA}$$

Все параметры модели значимы на 5%.

(a) Рассчитайте годовой доход мужчины, свободно владеющего английским, со стажем работы 2 года без сертификата CFA

(b) При прочих равных условиях на сколько отличаются годовые доходы мужчин и женщин в отрасли?

Задача 12. Исследователь оценил зависимость:
для 30 индустриально развитых стран:

$$\underset{(s.e.)}{\widehat{\ln \text{УА}}} = \underset{(0.86)}{-1.13} + \underset{(0.15)}{0.69} \ln \text{УН} - \underset{(0.16)}{0.64} \ln \text{УР}, \; RSS_1 = 27.63$$

для 30 развивающихся стран:

$$\underset{(s.e.)}{\widehat{\ln \text{УА}}} = \underset{(0.87)}{-1.12} + \underset{(0.13)}{0.81} \ln \text{УН} - \underset{(0.11)}{0.09} \ln \text{УР}, \; RSS_1 = 32.18$$

по всей выборке:

$$\underset{(s.e.)}{\widehat{\ln \text{УА}}} = \underset{(0.97)}{-1.13} + \underset{(0.21)}{0.75} \ln \text{УН} - \underset{(0.24)}{0.35} \ln \text{УР}, \; RSS_1 = 123.76$$

Проверьте гипотезу о том, что зависимость уровня активности в теневой экономике *УА* от уровня налогов *УН* и уровня правительственных расходов на борьбу с теневой экономикой *УР* одинаковая для развитых и развивающихся стран. Используйте уровень значимости 5%. Выпишите

(a) Нулевую и альтернативную гипотезы

(b) Тестовую статистику

(c) Распределение тестовой статистики

(d) Наблюдаемое значение тестовой статистики

(e) Область, в которой H_0 не отвергается

(f) Статистический вывод

Задача 13. Пусть регрессионная модель имеет вид $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$. Тестируется гипотеза $H_0 : \beta_1 = \beta_2 = 1$. Напишите уравнение для модели «с ограничением».

Задача 14. Пусть регрессионная модель имеет вид $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$. Тестируется гипотеза $H_0 : \begin{cases} \beta_1 + \beta_2 + \beta_3 = 0 \\ \beta_2 + \beta_3 = 0 \end{cases}$. Напишите уравнение для модели «с ограничением».

Задача 15*. При помощи метода наименьших квадратов оценивается модель $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ по 34 наблюдениям, $RSS = 15$, $TSS = 20$. На уровне значимости 5% проверьте гипотезу

$$H_0 : \begin{cases} \beta_1 + \beta_2 + \beta_3 = -3 \\ \beta_1 = -1 \\ \beta_2 = -1 \\ \beta_3 = -1 \end{cases}$$

если известно, что $\sum_{i=1}^{3} 4(y_i - \overline{y} + x_{i1} - \overline{x_1} + x_{i1} - \overline{x_1} + x_{i2} - \overline{x_2} + x_{i3} - \overline{x_3}) = 19$. Выпишите

(a) Тестовую статистику (формулу)

(b) Распределение тестовой статистики

(c) Наблюдаемое значение тестовой статистики

(d) Область, в которой H_0 не отвергается

(e) Статистический вывод

Задача 16. Пусть регрессионная модель $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, $i = 1, \ldots, n$, задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = \begin{bmatrix} \alpha & \beta_1 & \beta_2 \end{bmatrix}^T$. Известно, что $\varepsilon \sim N(0, \sigma^2 I)$ — n-мерный нормальный случайный вектор. Данные о наблюдениях y и X следующие:

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \ X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Для удобства расчётов ниже приведены матрицы:

$$X^T X = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ и } (X^T X)^{-1} = \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 3/2 \end{bmatrix}$$

(a) Укажите число наблюдений

(b) Укажите число регрессоров в модели (с учётом свободного члена)

(c) Рассчитайте $TSS = \sum_{i=1}^{n} (y_i - \overline{y})^2$

(d) Рассчитайте при помощи метода наименьших квадратов оценку для вектора неизвестных коэффициентов

(e) Найдите $RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

(f) Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оцененного уравнения регрессии

(g) Использую приведённые выше данные, рассчитайте несмещенную оценку для неизвестного параметра σ^2 регрессионной модели

(h) Использую приведённые выше данные, рассчитайте границы 80%-ого доверительного интервала для неизвестного параметра регрессии σ^2

(i) Использую приведённые выше данные, рассчитайте границы 80%-ого доверительного интервала для неизвестного параметра регрессии σ

(j) Использую имеющиеся данные, рассчитайте несмещенную оценку для ковариационной матрицы вектора МНК-оценок

(k) Использую приведённые выше данные, рассчитайте границы 90%-ого доверительного интервала для неизвестного параметра регрессии β_1

(l) Укажите «приближенно» вероятность того, что построенный в предыдущем пункте доверительный интервал не накроет истинное значение параметра β_1

(m) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + \beta_1 + \beta_2$

(n) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + \beta_1 - \beta_2$

(o) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + 2\beta_1 - 2\beta_2$

(p) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + 2\beta_1 - 3\beta_2$

Ответы.

(a) $n = 5$

(b) $k = 3$

(c) $TSS = 10$

(d) $\hat{\beta} = \begin{bmatrix} 2 & 1 & 2 \end{bmatrix}^T$

(e) $RSS = 4$

(f) $R^2 = 0.6$, качество регрессии - среднее

(g) $\hat{\sigma}^2 = 2$

(h) $[0.86, 18.98]$

(i) $[0.93, 4.35]$

(j) $\widehat{\mathrm{Cov}}(\hat{\beta}) = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$

(k) $[-3.12, 5.12]$

(l) 10%

(m) $[-1.08, 11.08]$

(n) $[-9.53, 11.53]$

(o) $[-21.51, 21.51]$

(p) $[-30.54, 26.54]$

Задача 17. Пусть регрессионная модель $y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, $i = 1, \ldots, n$, задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = \begin{bmatrix} \alpha & \beta_1 & \beta_2 \end{bmatrix}^T$. Известно, что $\varepsilon \sim N(0, \sigma^2 I)$ — n-мерный нормальный случайный вектор. Данные о наблюдениях y и X следующие:

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \; X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Для удобства расчётов ниже приведены матрицы:

$$X^T X = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ и } (X^T X)^{-1} = \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 3/2 \end{bmatrix}$$

(a) Укажите число наблюдений

(b) Укажите число регрессоров в модели (с учётом свободного члена)

(c) Рассчитайте $TSS = \sum_{i=1}^{n} (y_i - \overline{y})^2$

(d) Рассчитайте при помощи метода наименьших квадратов оценку для вектора неизвестных коэффициентов

(e) Найдите $RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

(f) Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оцененного уравнения регрессии

(g) Использую приведённые выше данные, рассчитайте несмещенную оценку для неизвестного параметра σ^2 регрессионной модели

(h) Использую приведённые выше данные, рассчитайте границы 80%-ого доверительного интервала для неизвестного параметра регрессии σ^2

(i) Использую приведённые выше данные, рассчитайте границы 80%-ого доверительного интервала для неизвестного параметра регрессии σ

(j) Использую имеющиеся данные, рассчитайте несмещенную оценку для ковариационной матрицы вектора МНК-оценок

(k) Использую приведённые выше данные, рассчитайте границы 90%-ого доверительного интервала для неизвестного параметра регрессии β_1

(l) Укажите «приближенно» вероятность того, что построенный в предыдущем пункте доверительный интервал не накроет истинное значение параметра β_1

(m) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + \beta_1 + \beta_2$

(n) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + \beta_1 - \beta_2$

(o) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + 2\beta_1 - 2\beta_2$

(p) Постройте 95%-ый доверительный интервал для следующей функции от неизвестных параметров $\alpha + 2\beta_1 - 3\beta_2$

Ответы.

(a) $n = 5$

(b) $k = 3$

(c) $TSS = 10$

(d) $\hat{\beta} = \begin{bmatrix} 2 & 3/2 & 1/2 \end{bmatrix}^T$

(e) $RSS = 6.5$

(f) $R^2 = 0.35$, качество регрессии - низкое

(g) $\hat{\sigma}^2 = 13/4$

(h) $[1.41, 30.84]$

(i) $[1.18, 5.55]$

(j) $\widehat{\mathrm{Cov}}(\hat{\beta}) = \begin{bmatrix} 13/8 & -13/8 & 0 \\ -13/8 & 13/4 & -13/8 \\ 0 & -13/8 & 39/8 \end{bmatrix}$

(k) $[-3.76, 6.76]$

(l) 10%

(m) $[-3.75, 11.75]$

(n) $[-10.43, 16.43]$

(o) $[-23.42, 31.42]$

(p) $[-32.88, 39.88]$

Задача 18. Оценивается регрессионная модель

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, i = 1, \ldots, n$$

где $\varepsilon_1, \ldots, \varepsilon_n$ — независимые случайные величины с математическим ожиданием 0 и дисперсией σ^2. Результаты оценивания приведены в следующей таблице

	Оценка	Ст. ошибка	t-статистика	P-значение
Константа	1.44	1.25		0.26
X1	0.82	0.09	9.10	0.00
X2			7.44	0.00

Также известно, что $TSS = 183.3939$, $RSS = 30.4118$, а число наблюдений $n = 25$. Кроме этого, 95%-ый доверительный интервал для $\hat{\beta}_2$ — $[0.3737, 0.6625]$. Найдите

(a) R^2 и R^2_{adj}

(b) Стандартную ошибку $\hat{\sigma}$

(c) ESS

(d) Значение F-статистики гипотезы о незначимости регрессии в целом

(e) Пропущенные в таблице значения,

Решение.

(a) $R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{30.4118}{183.3939} = 0.8342$
$R^2_{adj} = 1 - \frac{RSS/(n-k)}{TSS/(n-1)} = 1 - \frac{30.4118/(25-3)}{183.3939/(25-1)} = 0.8191$

(b) $\hat{\sigma} = \sqrt{\frac{RSS}{n-k}} = \sqrt{\frac{30.4118}{25-3}} = 1.3824$

(c) $ESS = TSS - RSS = 183.3939 - 30.4118 = 152.9821$

(d) $F = \frac{R^2}{1-R^2} \cdot \frac{n-k}{k} = \frac{0.8342}{1-0.8342} \cdot \frac{25-3}{3} = 55.3338$

(e) $t_{\hat{\alpha}} = \frac{\hat{\alpha}}{\hat{\sigma}_{\hat{\alpha}}} = \frac{1.4433}{1.2477} = 1.1567$

Заметим, что нижняя и верхняя границы 95%-ого доверительного интервала для неизвестного параметра β_2 равны соответственно:

$$\hat{\beta}_2 - t_{0.975,22} \cdot \hat{\sigma}_{\hat{\beta}_2} \text{ и } \hat{\beta}_2 + t_{0.975,22} \cdot \hat{\sigma}_{\hat{\beta}_2}$$

где $t_{0.975,22}$ означает квантиль уровня 0.975 для t-распределения с 22 степенями свободы[1]. Следовательно:

$$\hat{\beta}_2 - t_{0.975,22} \cdot \hat{\sigma}_{\hat{\beta}_2} = 0.3737 \text{ и } \hat{\beta}_2 + t_{0.975,22} \cdot \hat{\sigma}_{\hat{\beta}_2} = 0.6625$$

Складывая эти уравнения получаем:

$$2\hat{\beta}_2 = 1.0363$$

Значит,

$$\hat{\beta}_2 = 0.5181$$

Из формулы

$$t_{\hat{\beta}_2} = \frac{\hat{\beta}_2}{\hat{\sigma}_{\hat{\beta}_2}}$$

получаем:

$$\hat{\sigma}_{\hat{\beta}_2} = 0.0696$$

Задача 19. Оценивается регрессионная модель

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, i = 1, \dots, n$$

где $\varepsilon_1, \dots, \varepsilon_n$ — независимые случайные величины с математическим ожиданием 0 и дисперсией σ^2. Результаты оценивания приведены в следующей таблице

Также известно, что $TSS = 1370.9637$, $ESS = 1340.5519$, а число наблюдений $n = 25$. Кроме этого, 95%-ый доверительный интервал для $\hat{\beta}_2$ — $[1.379, 1.61]$. Найдите

[1] Другими словами, $t_{0.975,22}$ — это такая точка, что $\int_{-\infty}^{t_{0.975,22}} f_{t_{22}}(x)dx = 0.975$, где $f_{t_{22}}(x)$ — плотность t-распределения с 22 степенями свободы

	Оценка	Ст. ошибка	t-статистика	P-значение
Константа	0.91		1.18	0.25
X1	0.90	0.04		0.00
X2	1.49			0.00

(a) R^2

(b) Стандартную ошибку $\hat{\sigma}$

(c) RSS

(d) Значение F-статистики гипотезы о незначимости регрессии в целом

(e) Пропущенные в таблице значения,

Ответы.

(a) $R^2 = 0.9778$

(b) $\hat{\sigma} = 1.3824$

(c) $RSS = 30.4118$

(d) $F = 484.8796$

(e) $\hat{\sigma}_{\hat{\alpha}} = 0.7754$, $t_{\hat{\beta}_1} = 22.1836$, $\hat{\sigma}_{\hat{\beta}_2} = 0.0557$, $t_{\hat{\beta}_2} = 26.8343$

Задача 20. Оценивается регрессионная модель

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, i = 1, \ldots, n$$

где $\varepsilon_1, \ldots, \varepsilon_n$ — независимые случайные величины с математическим ожиданием 0 и дисперсией σ^2. Результаты оценивания приведены в следующей таблице

	Оценка	Ст. ошибка	t-статистика	P-значение
Константа		1.19	-1.46	0.16
X1	0.60		3.67	0.00
X2			7.44	0.00

Также известно, что $TSS = 111.5347$, $R^2 = 0.7273$, а число наблюдений $n = 25$. Кроме этого, 95%-ый доверительный интервал для $\hat{\beta}_2$ — $[0.7475, 1.325]$. Найдите

(a) Стандартную ошибку $\hat{\sigma}$

(b) ESS

(c) RSS

(d) Значение F-статистики гипотезы о незначимости регрессии в целом

(e) Пропущенные в таблице значения,

Ответы.

(a) $\hat{\sigma} = 1.3824$

(b) $ESS = 81.1228$

(c) $RSS = 30.4118$

(d) $F = 29.3422$

(e) $\hat{\alpha} = -1.7373$, $\hat{\sigma}_{\hat{\beta}_1} = 0.1621$, $\hat{\beta}_2 = 1.0363$, $\hat{\sigma}_{\hat{\beta}_2} = 0.1392$