Ф.И.О____

Группа__________________

ВОПРОСЫ

Пусть $Y_t = \beta_1 + \beta_2 X_{t2} + ... + \beta_k X_{tk} + \varepsilon_t \ \left(t = \overline{1,n}\right)$ - модель линейной регрессии.

Вопрос 1. Ошибки $\{\varepsilon_t\}_{t=1}^n$ называются ____________________________________, если существуют номера наблюдений t и s, такие, что $1 \le t \le n$; $1 \le s \le n$, для которых $\mathrm{D}(\varepsilon_t) \ne \mathrm{D}(\varepsilon_s)$.

Вопрос 2. Ошибки $\{\varepsilon_t\}_{t=1}^n$ называются ____________________________________, если существуют номера наблюдений t и s, такие, что $1 \le t \le n$; $1 \le s \le n$; $t \ne s$, для которых $\mathbb{E}(\varepsilon_t \varepsilon_s) \ne 0$

Вопрос 3. Переменная $X_j = \{X_{tj}\}_{t=1}^n$ называется ___________________________, если для всех $t = 1,...,n$ и $s = 1,...,n$ выполняется равенство $\operatorname{cov}\left(X_{tj}, \varepsilon_s\right) = 0$.

Вопрос 4. Статистика $\dfrac{\sum_{t=2}^{n}\left(\hat{\varepsilon}_t^2 - \hat{\varepsilon}_{t-1}^2\right)}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$ называется статистикой

____.

Вопрос 5. Статистика $\dfrac{\sum_{t=2}^{n}\left(\hat{\varepsilon}_t^2 - \hat{\varepsilon}_{t-1}^2\right)}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$ используется для тестирования

___________________ и не применяется, если в модели нет _______________________.

Вопрос 6. Оценки, которые являются решением задачи

$$\sum_{t=1}^{n}\left(Y_t - \beta_1 - \beta_2 X_{t2} - ... - \beta_k X_{tk}\right)^2 \to \min_{\beta_1,...,\beta_k},$$

называются ____.

Вопрос 7. Оценка $\left(X^T \mathrm{V}(\varepsilon)^{-1} X\right)^{-1} X^T \mathrm{V}(\varepsilon)^{-1} Y$ при $\mathrm{V}(\varepsilon) \ne \sigma^2 \cdot I$ используется для оценивания вектора неизвестных параметров β в случае, если ошибки $\{\varepsilon_t\}_{t=1}^n$ в модели являются ______________________________ или __________________________________.

Вопрос 8. Если зависимая переменная Y_t может принимать только два значения 0 и 1, то класс таких регрессионных моделей называется

____.

Вопрос 9. Пусть регрессионная модель имеет вид

$$\mathbb{P}\left(\{Y_t = 1\}\right) = \Phi\left(\beta_1 + \beta_2 X_{t2} + ... + \beta_k X_{tk}\right),$$

где $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$. Тогда такая модель называется ________________________.

Вопрос 10. Регрессионные модели вида $\mathbb{P}\left(\{Y_t = 1\}\right) = \Lambda\left(\beta_1 + \beta_2 X_{t2} + ... + \beta_k X_{tk}\right)$, где $\Lambda(x) = \dfrac{\exp(x)}{1 + \exp(x)}$, оцениваются, например, при помощи метода

____.

ЗАДАЧИ

Задача 1. 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 нормальные случайные величины с математическим ожиданием 0 и дисперсией σ^2. Результаты оценивания приведены в следующей таблице.

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	0,83
R-квадрат	**???**
Нормированный R-квадрат	0,66
Стандартная ошибка	**???**
Наблюдения	25,00

Дисперсионный анализ

	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	2,00	96,94	48,47	???	0,00
Остаток	22,00	**???**	1,96		
Итого	24,00	140,04			

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
Y-пересечение	7,34	**???**	8,50	0,00	5,55	9,13
X1	0,78	0,13	**???**	0,00	0,51	1,06
X2	0,75	**???**	**???**	0,00	0,44	1,06

Заполните следующую таблицу.

A.	$R^2 =$	
B.	$\widehat{\sigma} =$	
C.	$RSS =$	
D.	$F =$	
E.	$\widehat{\sigma}_{\hat{\alpha}} =$	
F.	$t_{\widehat{\beta}_1} =$	
G.	$\widehat{\sigma}_{\widehat{\beta}_2} =$	
H.	$t_{\widehat{\beta}_2} =$	

Задача 2. Была оценена функция Кобба-Дугласа с учетом человеческого капитала H (K – физический капитал, L – труд)

$$\widehat{\ln(Q)} = 1.4 + 0.46\ln(L) + 0.27\ln(H) + 0.23\ln(K)$$

$$ESS = 170.4, \quad RSS = 80.3, \quad n = 21.$$

A. Чему равен коэффициент R^2_{adj}.

B. Проверьте регрессию на значимость “в целом” на уровне значимости 5%.

Заполните следующую таблицу.

A.	$R^2_{adj} =$
B0.	H_0 : H_1 :
B1.	*тестовая статистика*:
B2.	*распределение тестовой статистики*:
B3.	*наблюдаемое значение тестовой статистики*:
B4.	*область, в которой H_0 не отвергается*:
B5.	*статистический вывод*:

Задача 3. Рассмотрим следующую модель зависимости цены дома $Price$ (в тысячах долларов) от его площади $Hsize$ (в м^2), площади участка $Lsize$ (в м^2), числа ванных комнат $Bath$ и числа спален BDR:

$$\widehat{Price} = \widehat{\alpha} + \widehat{\beta_1} \cdot Hsize + \widehat{\beta_2} \cdot Lsize + \widehat{\beta_3} \cdot Bath + \widehat{\beta_4} \cdot BDR,$$

$$R^2 = 0.218, \quad n = 23.$$

Напишите спецификацию регрессии с ограничениями для проверки статистической гипотезы $H_0 : \beta_3 = 20\beta_4$. Для регрессии с ограничениями был вычислен коэффициент $R_R^2 = 0.136$. Протестируйте нулевую гипотезу на уровне значимости 5%.

Заполните таблицу.

A.	*спецификация*:
B0.	H_0: H_1:
B1.	*тестовая статистика*:
B2.	*распределение тестовой статистики*:
B3.	*наблюдаемое значение тестовой статистики*:
B4.	*область, в которой* H_0 *не отвергается*:
B5.	*статистический вывод*:

Задача 4. Пусть регрессионная модель $Y_i = \alpha + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \varepsilon_i \ \left(i = \overline{1,n}\right)$ задана в матричном виде при помощи уравнения $Y = X\beta + \varepsilon$, где $\beta = \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix}$. Известно, что $\mathbb{E}\left(\varepsilon\right) = \mathbf{0}$ и $\mathrm{V}\left(\varepsilon\right) = \sigma^2 \cdot I$. Известно также, что

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

Для удобства расчетов ниже приведены матрицы

$$X^T X = \begin{bmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ и } \left(X^T X\right)^{-1} = \begin{bmatrix} 0.5 & -0.5 & 0.0 \\ -0.5 & 1.0 & -0.5 \\ 0.0 & -0.5 & 1.5 \end{bmatrix}.$$

Выполните следующие задания и заполните таблицу ниже.

A. Рассчитайте $TSS = \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2$.

B. Рассчитайте при помощи метода наименьших квадратов оценку для вектора неизвестных коэффициентов.

C. Чему равен $\widehat{\varepsilon_5}$ – МНК-остаток регрессии, который соответствует 5-му наблюдению?

D. Найдите $RSS = \sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2$.

E. Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оцененного уравнения регрессии.

F. Используя приведенные выше данные, рассчитайте несмещенную оценку для неизвестного параметра σ^2 регрессионной модели.

G. Рассчитайте $\widehat{\mathrm{V}}\left(\widehat{\beta}\right)$ – оценку для ковариационной матрицы вектора МНК-коэффициентов $\widehat{\beta}$.

H. Найдите $\widehat{\mathrm{D}}\left(\widehat{\alpha}\right)$ – несмещенную оценку дисперсии МНК-коэффициента $\widehat{\alpha}$.

I. Найдите $\widehat{\mathrm{D}}\left(\widehat{\beta_1}\right)$ – несмещенную оценку дисперсии МНК-коэффициента $\widehat{\beta_1}$.

J. Найдите $\widehat{\text{cov}}\left(\hat{\alpha}, \widehat{\beta_1}\right)$ – несмещенную оценку ковариации МНК-коэффициентов $\hat{\alpha}$ и $\widehat{\beta_1}$.

K. Найдите $\widehat{\text{corr}}\left(\hat{\alpha}, \widehat{\beta_1}\right)$ – оценку ковариации МНК-коэффициентов $\hat{\alpha}$ и $\widehat{\beta_1}$.

L. Найдите $\widehat{\text{D}}\left(\hat{\alpha} - \widehat{\beta_1}\right)$.

A.	
B.	
C.	
D.	
E.	
F.	
G.	
H.	
I.	
J.	
K.	
L.	