Часть 1. Тест.

Вопрос 1 ♣ Если модули остатков оценённой с помощью МНК регрессионной модели линейно и значимо зависят от квадрата регрессора Z, то гетероскедастичность можно попытаться устранить,

- A умножив исходное уравнение на Z
- B умножив исходное уравнение на Z^2
- C поделив исходное уравнение на $\sqrt{Z}$
- D умножив исходное уравнение на $\sqrt{Z}$
- E поделив исходное уравнение на Z^2
- F поделив исходное уравнение на Z
- G *Нет верного ответа.*

Вопрос 2 ♣ Выберите верное утверждение про метод максимального правдоподобия применительно к линейной регрессии.

- A ММП не может быть применим при нестрогой мультиколлинеарности
- B ММП может быть применим только при нормально распределённых ошибках
- C ММП может быть применим только в случае гетероскедастичности
- D ММП требует спецификации семейства распределения для ошибок
- E ММП может быть применим только в случае гомоскедастичности
- F *Нет верного ответа.*

Вопрос 3 ♣ При наличии условной гетероскедастичности нарушается следующее свойство МНК-оценок параметров классической регрессии:

- A эффективность в классе линейных и несмещенных оценок
- B линейность по зависимой переменой
- C несмещённость
- D равенство нулю суммы остатков
- E ортогональность вектора остатков и вектора прогнозов
- F *Нет верного ответа.*

Вопрос 4 ♣ Оценка максимального правдоподобия параметра λ по случайной выборке X_1, ..., X_n из распределения с функцией плотности

$$f(x|\lambda) = \begin{cases} (2\lambda+1)\exp(-x)\exp(-2\lambda x), \text{ если } 0 < x; \\ 0, \text{ иначе.} \end{cases}$$

имеет вид:

- A $\hat{\lambda}_{ML} = \frac{1}{2\bar{X}} - \frac{1}{2}$
- B $\hat{\lambda}_{ML} = \frac{1}{2\bar{X}} + 1$
- C $\hat{\lambda}_{ML} = \bar{X} - \frac{1}{2}$
- D $\hat{\lambda}_{ML} = \frac{1}{\bar{X}} + \frac{1}{2}$
- E $\hat{\lambda}_{ML} = \frac{1}{2\bar{X}} + \frac{1}{2}$
- F *Нет верного ответа.*

Вопрос 5 ♣ Методом максимального правдоподобия Гоша оценил модель

$$Y_i = \beta_1 + \beta_2 X_{i2} + \ldots + \beta_6 X_{i6} + \varepsilon_i,$$

где $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 I)$, по 12 наблюдениям. Оказалось, что $RSS = 24$. Оценка дисперсии случайной составляющей равна

- A 2.4
- B 24/7
- C 2
- D 0.5
- E 0.48
- F *Нет верного ответа.*

Вопрос 6 ♣ Имеются данные по 100 работникам: затраты на проезд в общественном транспорте (E_i, руб.), количество часов работы в день (WH_i, руб.), количество часов отдыха в день (LH_i, руб.) и количество часов сна в день (SH_i, руб.). Считая, что всё время суток распределяется между трудом, сном и отдыхом, оценка регрессии в виде

$$E_i = \beta_1 + \beta_2 WH_i + \beta_3 LH_i + \beta_4 SH_i + u_i$$

приведет к тому, что

- A МНК–оценки получить не удастся
- B МНК-оценки параметров окажутся смещёнными
- C МНК-оценки параметров окажутся неэффективными в классе линейных и несмещённых
- D коэффициент детерминации R^2 окажется отрицательным
- E МНК-оценки параметров регрессии будут несмещенными и эффективными
- F *Нет верного ответа.*

Вопрос 7 ♣ Взвешенный МНК служит для оценивания регрессионных моделей в случае нарушений следующего условия теоремы Гаусса-Маркова:

- A Величина Y_i линейна по β_1, β_2, ...
- B $\mathrm{Var}(u_i) = \sigma^2$
- C $\operatorname{rank} X = k$
- D $\mathbb{E}(u_i) = 0$
- E u_i распределены нормально
- F *Нет верного ответа.*

Вопрос 8 ♣ По $n = 650$ наблюдениям была оценена регрессия:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + u_i.$$

Затем была оценена регрессия $|\hat{u}_i| = \alpha_1 + \alpha_2 \frac{1}{Z_i} + \nu_i$. Оказалось, что $\hat{\alpha}_2 = 36$ и $se(\hat{\alpha}_2) = 5$. Согласно этим данным, на уровне значимости 5% гипотеза о

- A верной функциональной форме не отвергается
- B пропущенной переменной $1/Z_i$ не отвергается
- C гомоскедастичности не отвергается
- D пропущенной переменной $1/Z_i$ отвергается
- E гомоскедастичности отвергается
- F верной функциональной форме отвергается
- G *Нет верного ответа.*

Вопрос 9 ♣ Василий хочет оценить константу μ в модели $Y_i = \mu + u_i$, где $\mathbb{E}(u_i) = 0, \mathbb{E}(u_i u_j) = 0$ при $i \neq j$, $\mathrm{Var}(u_i) = \sigma^2 X_i$ и $X_i > 0$.
В классе линейных несмещенных оценок наиболее эффективной является:

- A $\frac{\sum Y_i/\sqrt{X_i}}{\sum 1/X_i}$
- B $\bar{Y}$
- C $\frac{\sum Y_i/X_i}{\sum 1/X_i^2}$
- D $\frac{\sum Y_i X_i}{\sum X_i^2}$
- E $\frac{\sum Y_i X_i}{\sum X_i}$
- F $(I'I)^{-1}I'Y$
- G *Нет верного ответа.*

Вопрос 10 ♣ Василий оценивает модель $Y_i = \beta_1 + \beta_2 X_i + u_i$ с помощью МНК при двух разных предположениях. Предположение А: $u_i \sim \mathcal{N}(0; \sigma^2)$. Предположение В: u_i распределены экспоненциально с параметром λ. При этом окажется, что

- A $\hat{\beta}_1^A > \hat{\beta}_1^B, \hat{\beta}_2^A = \hat{\beta}_2^B$
- B $\hat{\beta}_1^A = \hat{\beta}_1^B, \hat{\beta}_2^A = \hat{\beta}_2^B$
- C $\hat{\beta}_1^A = \hat{\beta}_1^B, \hat{\beta}_2^A < \hat{\beta}_2^B$
- D $\hat{\beta}_1^A < \hat{\beta}_1^B, \hat{\beta}_2^A = \hat{\beta}_2^B$
- E $\hat{\beta}_1^A = \hat{\beta}_1^B, \hat{\beta}_2^A > \hat{\beta}_2^B$
- F *Нет верного ответа.*

Вопрос 1 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$ $\boxed{\text{G}}$

Вопрос 2 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 3 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 4 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 5 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 6 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 7 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Вопрос 8 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$ $\boxed{\text{G}}$

Вопрос 9 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$ $\boxed{\text{G}}$

Вопрос 10 : $\boxed{\text{A}}$ $\boxed{\text{B}}$ $\boxed{\text{C}}$ $\boxed{\text{D}}$ $\boxed{\text{E}}$ $\boxed{\text{F}}$

Часть 2. Задачи.

1. По 35 наблюдениям сотрудники НИИ оценили уравнение регрессии $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ и рассчитали остатки ε_i. После того они приступили к диагностике возможных недостатков модели, обнаружили гетероскедастичность и решили её побороть.

   а) Самый младший научный сотрудник выдвинул предположение, что стандартное отклонение случайной составляющей может быть выражено так: $\sigma_{\varepsilon,i} = ax_i$, где a — неизвестный коэффициент. Каким образом нужно преобразовать исходное уравнение регрессии, чтобы избавиться от гетероскедастичности?

   б) Профессор решил перепроверить результаты и оценил регрессию:

   $$\hat{e}_i^2 = -0.3 + 0.08x_i - 0.01x_i^2, R^2 = 0.15$$

   Свидетельствует ли полученный профессором результат о наличии гетероскедастичности?

2. Были обследованы 36 предприятий по трём показателям: K_i — основным фондам (млн. руб.), W_i — фонду оплаты труда (млн. руб.), R_i — расходам на НИОКР (млн. руб.). Получены оценки вектора средних $\hat{\mu} = (3, 4, 5)'$ и ковариационной матрицы $\hat{\Sigma} = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 11 & 0 \\ 3 & 0 & 10 \end{pmatrix}$.

   Предполагая, что приведение переменных к общему масштабу не требуется, найдите первую главную компоненту и определите долю суммарной дисперсии, которую она объясняет.

3. Используя 80 наблюдений, исследователь оценил две конкурирующие модели: $\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z$, в которой $RSS_1 = 36875$ и $\widehat{\ln y} = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z$, в которой $RSS_2 = 122$.

   Выполнив преобразование $y_i^* = y_i / \sqrt[n]{\prod y_i}$, исследователь также оценил две вспомогательные регрессии: $\hat{y}^* = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z$, в которой $RSS_1^* = 239$ и $\widehat{\ln y^*} = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z$, в которой $RSS_2^* = 121$.

   а) Проинтерпретируйте коэффициент $\hat{\beta}_2$ в двух конкурирующих моделях

   б) С помощью подходящего теста выберите наилучшую из двух конкурирующих моделей