Имя, фамилия и номер группы:

. . . .

1. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
2. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
3. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
4. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
5. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
6. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
7. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
8. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
9. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
10. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
11. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
12. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
13. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
14. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
15. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
16. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
17. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
18. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
19. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
20. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
21. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
22. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
23. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
24. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
25. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
26. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
27. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
28. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
29. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f
30. ☐ a ☐ b ☐ c ☐ d ☐ e ☐ f

Удачи!

1. Какая зависимость математического ожидания исходного процесса от времени предполагается в альтернативной гипотезе ADF-теста с константой?

   a) нет верного ответа
   b) $\mathbb{E}(y_t)$ строго возрастает
   c) $\mathbb{E}(y_t) = \mu + \alpha t + \beta t^2$
   d) $\mathbb{E}(y_t)$ строго монотонна
   e) $\mathbb{E}(y_t) = \mu + \alpha t$
   f) $\mathbb{E}(y_t) = \mu$

2. У меня набор данных из трёх наблюдений: $x = (1, 2, 3)$, $y = (5, 6, 10)$.
   Чему равна МНК-оценка $\hat{\beta}$ в модели $y_i = \beta + u_i$?

3. Леонардо оценил логит-регрессию $\hat{\mathbb{P}}(y_i = 1) = \Lambda(0.3 - 0.6x_i + 0.5z_i)$. Для некоторого наблюдения прогноз вероятности равен $\hat{\mathbb{P}}(y_i = 1) = 0.7$.
   Оцените предельный эффект увеличения z_i для этого наблюдения.

   a) 0.105
   b) недостаточно информации
   c) 0.5
   d) 0.15
   e) 0.35
   f) 0.7

4. Метод максимального правдоподобия для оценки коэффициентов регрессии $Y = X\beta + u$ НЕ МОЖЕТ быть применён, если

   a) $u \sim \mathcal{N}(0; \Omega)$ и $\Omega = 2017 \cdot I$, где I — единичная матрица
   b) $u \sim \mathcal{N}(0; \Omega)$ и структура Ω неизвестна
   c) $u \sim \mathcal{N}(0; \Omega)$ и $\Omega = b \cdot I$, где b — неизвестный параметр
   d) $u \sim \mathcal{N}(0; \Omega)$ и структура Ω известна, но зависит от набора неизвестных параметров
   e) закон распределения вектора u известен, но не является нормальным

5. Чудо-швабры производятся на разных заводах по одной из двух технологий, A или B. Исследователь оценил две модели зависимости выпуска, Y, от количества сырья, X, и технологии:
   $\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 A_i + \hat{\alpha}_2 X_i + \hat{\alpha}_3 A_i X_i$;
   $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 B_i + \hat{\beta}_2 X_i + \hat{\beta}_3 B_i X_i$.
   Переменная A_i равна единице для заводов с технологией A и нулю иначе, а переменная B_i равна единице для заводов с технологией B и нулю иначе.
   Оценки коэффициентов связаны соотношением

   a) $\hat{\alpha}_1 = \hat{\beta}_0$
   b) $\hat{\alpha}_0 = \hat{\beta}_0$
   c) $\hat{\alpha}_2 = \hat{\beta}_2$
   d) $\hat{\alpha}_0 + \hat{\alpha}_1 = \hat{\beta}_0$
   e) $\hat{\alpha}_0 = \hat{\beta}_0 + \hat{\beta}_1$

6. Рассмотрим алгоритм градиентного бустинга над решающими деревьями для задачи регрессии. Каждое дерево даёт свой прогноз.
   Выберите верное утверждение о механизме агрегирования прогнозов отдельных деревьев в прогноз всего алгоритма.

   a) прогнозы деревьев складываются и домножаются на темп обучения
   b) вес, с которым прогноз дерева учитывается в итоговом прогнозе, нелинейно убывает по номеру дерева
   c) вес, с которым прогноз дерева учитывается в итоговом прогнозе, линейно убывает по номеру дерева
   d) вес, с которым прогноз дерева учитывается в итоговом прогнозе, линейно растёт по номеру дерева
   e) нет верного ответа
   f) вес, с которым прогноз дерева учитывается в итоговом прогнозе, нелинейно растёт по номеру дерева

7. При оценивании модели $Y_t = X_t'\beta + u_t$ была обнаружена автокорреляция первого порядка с $\hat{\rho} = -0.6$. Чтобы провести корректное оценивание, можно применить метод наименьших квадратов к преобразованным данным. При этом первое наблюдение окажется домноженным на

   a) $\sqrt{0.6}$
   b) 0.8
   c) $\sqrt{0.84}$
   d) 0.4
   e) 0.6
   f) -0.6

8. Чебурашка оценил модель $Y_i = \beta_0 + \beta_1 X_i + u_i$, а Крокодил Гена — модель $X_i = \gamma_0 + \gamma_1 Y_i + u_i$. Оказалось, что $\hat{\gamma}_1 = 0.25/\hat{\beta}_1$. Величина R^2 в регрессии Чебурашки равна

   a) 0.5
   b) 0.25
   c) 0.75
   d) 1
   e) 0

9. Рассмотрим модель
$$y_i = \beta_1 + \beta_2 x_i + u_i.$$
Буквой X обозначена матрица всех предикторов. Выполнены классические предпосылки на ошибку $\mathbb{E}(u_i \mid X) = 0$, $\mathrm{Var}(u_i \mid X) = \sigma^2$, $\mathrm{Cov}(u_i \mid X) = 0$.
По 100 наблюдениям оказалось, что $\sum x_i^2 = 40$, $\sum (x_i - \bar{x})^2 = 20$.
Найдите величину $\mathrm{Var}(\hat{\beta}_2 \mid X)/\sigma^2$.
Ответ вводите с точность до двух знаков после десятичной точки.

10. Рассмотрим логистическую модель с константой и одним регрессором x_i. Обозначим логарифм шансов как $h(x_i) = \ln(\mathbb{P}(y_i = 1)/\mathbb{P}(y_i = 0))$.
Найдите верное утверждение про производную $h'(x_i)$.

   a) Производная $h'(x_i)$ линейна по x_i.
   b) Производная $h'(x_i)$ равна логистической функции от x_i.
   c) Производная $h'(x_i)$ равна константе.
   d) Производная $h'(x_i)$ достигает максимума только при $x_i = 0$.
   e) Производная $h'(x_i)$ произвольным образом зависит от x_i.
   f) Производная $h'(x_i)$ достигает минимума только при $x_i = 0$.

11. Известно, что регрессоры X и Z ортогональны, а истинная зависимость описывается уравнением $Y_i = \alpha_1 + \alpha_2 X_i + \alpha_3 Z_i + u_i$. Исследователь оценивает с помощью МНК две регрессии: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ и $\hat{Y}_i = \hat{\gamma}_1 + \hat{\gamma}_2 Z_i$. При этом

   a) $\hat{\beta}_2$ — смещённая оценка для α_2; $\hat{\gamma}_2$ — несмещённая оценка для α_3
   b) $\hat{\beta}_2$ — несмещённая оценка для α_2; $\hat{\gamma}_2$ — смещённая оценка для α_3
   c) $\hat{\beta}_2$ — несмещённая оценка для α_2; $\hat{\gamma}_2$ — несмещённая оценка для α_3
   d) $\hat{\beta}_2$ — смещённая оценка для α_2; $\hat{\gamma}_2$ — смещённая оценка для α_3
   e) $\hat{\beta}_2$ — эффективная оценка для α_2; $\hat{\gamma}_2$ — эффективная оценка для α_3

12. При оценивании коэффициентов моделей бинарного выбора

   a) оценки логит и пробит моделей имеют противоположные знаки
   b) оценки логит моделей всегда выше, чем пробит
   c) оценки логит и пробит моделей всегда совпадают
   d) оценки пробит моделей всегда выше, чем логит
   e) нет верного ответа
   f) оценки пробит модели имеют более высокую значимосить, чем логит

13. Модель коррекции ошибками имеет следующий вид

    a) $\Delta Y_t = \delta - \gamma(Y_{t-1} - \alpha - \beta X_{t-1}) + u_t$
    b) $\Delta Y_t = \delta + \phi\Delta X_{t-1} - \gamma(Y_{t-1}) + u_t$
    c) $\Delta Y_t = \delta + \phi\Delta X_{t-1} - \gamma(Y_{t-1} - \alpha - \beta X_{t-1}) + u_t$
    d) $Y_t = \delta + \phi\Delta X_{t-1} - \gamma(Y_{t-1} - \alpha - \beta X_{t-1}) + u_t$
    e) $Y_t = \delta - \gamma(Y_{t-1} - \alpha - \beta X_{t-1}) + u_t$

14. В линейной модели $Y_i = \beta_0 + \beta_1 X_i + u_i$ стохастический регрессор и случайный член u_i коррелированы. Состоятельные оценки коэффициентов можно получить с помощью

    a) метода главных компонент
    b) обобщённого МНК
    c) метода инструментальных переменных
    d) взвешенного МНК
    e) метода наименьших квадратов

15. Пантелеймон предполагает, что зависимость y от x линейная. У Пантелеймона данные по 10 городам. Он считает, что коэффициент наклона не зависит от города, а вот константа в каждом городе может быть своя.
    Всего у Пантелеймона 200 наблюдений.
    Сколько дамми-переменных понадобиться ввести Пантелеймону в дополнение к модели парной регрессии?

16. Если квадраты остатков оценённой с помощью МНК регрессионной модели линейно и значимо зависят от квадрата регрессора Z, то гетероскедастичность можно попытаться устранить,

    a) поделив исходное уравнение на Z
    b) поделив исходное уравнение на $\sqrt{Z}$
    c) поделив исходное уравнение на Z^2
    d) умножив исходное уравнение на Z^2
    e) умножив исходное уравнение на Z
    f) умножив исходное уравнение на $\sqrt{Z}$

17. Какой период у функции $\sin(5\pi t/365)$?
    Ответ укажите с точностью до двух знаков после десятичной точки.

18. Рассмотрим уравнение
    $$y_t = 14y_{t-1} + 19y_{t-2} + u_t,$$
    где (u_t) — белый шум.
    При каких начальных условиях получится нестационарное решение?

    a) только при $y_0 = 1$ и $y_1 = 1$
    b) нет верного ответа
    c) только при $y_0 = 0$ и $y_1 = 1$
    d) только при $y_0 = u_0$ и $y_1 = u_1$
    e) только при $y_0 = 0$ и $y_1 = 0$
    f) только при $y_0 = 1$ и $y_1 = 0$

19. В рамках ETS(AAN) модели постройте 95%-й предиктивный интервал на один шаг вперёд, Последнее значение сглаженного ряда (тренда) равно 100, последний наклон сглаженного ряда (тренда) равен 5, $\alpha = \beta = 0.3$, $u_t \sim \mathcal{N}(0; 16)$.
    В ответ введите правую границу предиктивного интервала с точностью до двух знаков после десятичной точки.

20. Крокодил Гена оценил с помощью МНК зависимость $Y_i = \beta_0 + \beta_1 X_i + u_i$. Оказалось, что $\hat\beta_0 = 90$, а $\hat\beta_1 = 3$. Чебурашка увеличил переменные X и Y на 10% и снова оценил уравнение регрессии. В результате этой корректировки

   a) оценки $\hat\beta_0$ и $\hat\beta_1$ увеличились
   b) оценка $\hat\beta_0$ увеличилась, а оценка $\hat\beta_1$ не изменилась
   c) оценка $\hat\beta_0$ уменьшилась, а оценка $\hat\beta_1$ не изменилась
   d) оценки $\hat\beta_0$ и $\hat\beta_1$ уменьшились
   e) оценки $\hat\beta_0$ и $\hat\beta_1$ не изменились

21. Крокодил Гена оценивает модель регрессии $Y_i = \beta_0 + \beta_1 X_i + u_i$ с помощью МНК. Чебурашка получит такую же оценку коэффициента β_1, если будет минимизировать

   a) выборочную дисперсию объясняемой переменной
   b) выборочную дисперсию объясняющей переменной
   c) выборочную дисперсию остатков
   d) коэффициент детерминации
   e) выборочную ковариацию регрессора и объясняемой переменной

22. Джеймс Бонд оценил парную регрессию и оказалось, что $\hat Y_i = 5 + 6X_i$. Если Джеймс Бонд оценит регрессию без константы, то окажется, что

   a) $\hat Y_i = 6X_i$
   b) нет верного ответа
   c) $\hat Y_i = 5.5X_i$
   d) $\hat Y_i = 11X_i$
   e) $\hat Y_i = 5.5$
   f) $\hat Y_i = 11$
   g) $\hat Y_i = 5$

23. Известно, что $y_t = 5 + u_t + 9u_{t-1}$, где (u_t) — белый шум с дисперсией 8.
Найдите частную автокорреляцию второго порядка ϕ_{22} процесса (y_t) с точностью до двух знаков после десятичной точки.

24. Рассмотрим логит-модель $\hat Y_i^* = \hat\beta_1 + \hat\beta_2 X_i + \hat\beta_3 D_i$, и $Y_i = 1$, если $Y_i^* > 0$. Если переменная X_i является количественной, то предельный эффект увеличения X_i можно посчитать по формуле

   a) $\hat\beta_2 / f^2(\hat Y_i^*)$
   b) $\hat\beta_2 / F(\hat Y_i^*)$
   c) $\hat\beta_2 f(\hat Y_i^*)$
   d) $\hat\beta_2 / F^2(\hat Y_i^*)$
   e) $\hat\beta_2 / f(\hat Y_i^*)$

25. Выберите верное утверждение про стационарный $ARMA(p, q)$ процесс относительно белого шума (u_t).

   a) является частным случаем $AR(p)$ процесса относительно (u_t).
   b) является частным случаем $MA(\infty)$ процесса относительно (u_t).
   c) является частным случаем $AR(q)$ процесса относительно (u_t).
   d) является частным случаем $MA(p)$ процесса относительно (u_t).
   e) является частным случаем $MA(q)$ процесса относительно (u_t).
   f) нет верного ответа

26. Из откровений внеземного разума известно, что эндогенности в модели $Y_i = \beta_0 + \beta_1 X_i + u_i$ нет. Однако Вовочка нашёл хороший инструмент z_i, отвечающий всем требованиям, предъявляемым к инструментам, и оценил β_1 методом инструментальных переменных. Его оценка β_1 окажется

   a) несостоятельной
   b) состоятельной, но не эффективной
   c) состоятельной и эффективной
   d) состоятельной
   e) невозможно сказать по имеющимся данным

27. Известны 95%-ые доверительные интервалы для коэффициентов регрессии: $\beta_1 \in [-4; 10]$, $\beta_2 \in [2; 10]$. На уровне значимости 5%

   a) $\hat{\beta}_1$ значим, $\hat{\beta}_2$ значим
   b) $\hat{\beta}_1$ значим, $\hat{\beta}_2$ не значим
   c) $\hat{\beta}_1$ не значим, $\hat{\beta}_2$ не значим
   d) $\hat{\beta}_1$ не значим, $\hat{\beta}_2$ значим
   e) значимость проверить невозможно

28. Условием теоремы Гаусса-Маркова, необходимым для несмещённости оценок коэффициентов регрессии в модели $Y_i = \beta_1 + \beta_2 X_i + u_i$ является

   a) некоррелированность случайных ошибок
   b) гомоскедастичность случайных ошибок
   c) гетероскедастичность случайных ошибок
   d) $\mathbb{E}(u_i) = 0$
   e) нормальность случайных ошибок
   f) $\mathbb{E}(u_i) \neq 0$

29. Найдите наименьшее возможное значение p, если известно, что $SARIMA(2,0,2)(3,0,3)[12]$-модель является частным случаем $ARMA(p,q)$-модели.

30. Оценка регрессионной зависимости с помощью МНК по 1234 наблюдениям имеет вид $\hat{Y}_i = 1 - 3X_i + 4Z_i$. Оценка ковариационной матрицы имеет вид

$$\operatorname{Var}(\hat{\beta}) = \begin{pmatrix} 1 & 0.1 & 0.2 \\ 0.1 & 4 & 1.5 \\ 0.2 & 1.5 & 18 \end{pmatrix}.$$

Длина 95%-го доверительного интервала для $\beta_2 + \beta_3$ примерно равна

   a) 5
   b) 25
   c) 10
   d) 1.96
   e) 20

31. По 546 наблюдениям за 1987 г. оценили зависимость стоимости частных домов в Канаде price (измеряемой в долларах США) от общей площади square (измеряемой в кв. м.) наличия подъездного пути driveway (1 — если есть, 0 — если нет):

переменная	коэффициент	ст. ошибка	t-статистика	P-значение
square	2.724	2.15	1.27	0.206
driveway	-922.563	8602.312	-0.11	0.915
square*driveway	3.479	1.43	2.42	0.016
$const$	38731.07	8156.39	4.75	0.000

Согласно полученным результатам, при уровне значимости 5%, наличие подъездного пути увеличивает стоимость каждого квадратного метра жилья на

a) 6.203 \$
b) 3.479 \$
c) 2.724 \$
d) -922.563 \$
e) 0

32. Рассмотрим модель

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i.$$

Оценка ковариационной матрицы коэффициентов равна

$$\begin{pmatrix} 10 & -2 & -1 \\ ? & 20 & 2 \\ ? & ? & 30 \end{pmatrix}$$

Найдите оценку дисперсии $\widehat{\mathrm{Var}}(\hat{\beta}_2 \mid X)$.

33. Храбрый исследователь Вениамин оценил регрессию $\hat{Y}_i = \underset{(5)}{23} + \underset{(2)}{10} X_i$, в скобках приведены стандартные ошибки. Доверительный интервал для свободного члена равен $[14; 32]$. Доверительный интервал для коэффициента наклона при том же уровне доверия будет равен

a) $[6.4; 13.6]$
b) $[6; 14]$
c) $[6.08; 13.92]$
d) $[1; 19]$
e) $[5; 15]$

34. У Винни-Пуха был временной ряд из 200 наблюдений. Винни использует алгоритм градиентного бустинга над деревьями, причем в качестве предикторов он берёт лаги y_{t-1}, ..., y_{t-8}.
По скольки наблюдениям обучается каждое из деревьев?

35. Стьюдентизированные остатки регрессии используются

a) в методе главных компонент
b) в тесте Саргана
c) на первом шаге двухшагового МНК
d) для выявления выбросов
e) на первом шаге при проведении теста Годфельда-Квандта

36. Рассмотрим веса, с которыми исходные переменные входят в первую главную компоненту. Выберите верное утверждение.

a) Веса могут быть абсолютно произвольными.
b) Сумма квадратов весов равна 1.
c) Сумма весов равна 0.
d) Сумма весов равна 1, веса могут быть отрицательными.
e) Сумма модулей весов равна 1.
f) Сумма весов равна 1, веса неотрицательные.

37. У Агнессы 120 наблюдений и она оценивает качество прогнозов на два шага вперёд с помощью кросс-валидации растущим окном. Стартовая ширина окна равна 70.
Сколько раз будет оценена каждая модель при выполнении кросс-валидации?

38. У $ARMA(2,3)$-модели с нулевым ожиданием и нормально распределенными ошибками логарифм правдоподобия оказался равен -130.
Найдите значение информационного критерия Акаике.

39. Аркадий оценил модель множественной регрессии с константой, а Борислав добавил в эту модель ещё два предиктора.
Они, соответственно, получили суммы квадратов остатков RSS_a и RSS_b, общие суммы квадратов TSS_a и TSS_b, и оценённые суммы квадратов ESS_a, ESS_b.
Выберите верное утверждение.

a) $TSS_a = TSS_b$, $RSS_a \geq RSS_b$.
b) $TSS_a = TSS_b$, $RSS_a \leq RSS_b$.
c) $ESS_a = ESS_b$, $TSS_a \leq TSS_b$.
d) $RSS_a = RSS_b$, $ESS_a \geq ESS_b$.
e) $ESS_a = ESS_b$, $TSS_a \geq TSS_b$.
f) $RSS_a = RSS_b$, $ESS_a \leq ESS_b$.

40. Гипотеза о неадекватности множественной регрессии проверяется с помощью статистики равной

a) $\frac{RSS}{TSS}$
b) $\frac{TSS/(n-1)}{ESS/(k-1)}$
c) $\frac{ESS/(k-1)}{RSS/(n-k)}$
d) $\frac{TSS/(n-1)}{RSS/(n-k)}$
e) $\frac{\hat{\beta}-\beta}{se(\hat{\beta})}$
f) $\frac{ESS}{TSS}$

41. Рассмотрим короткий ряд из четырех наблюдений: 5, 4, 6, 5.
Предположим, что ряд описывается моделью $y_t = \mu + u_t$, где величины (u_t) независимы и нормально распределены $\mathcal{N}(0;\sigma^2)$.
Найдите оценку $\hat{\sigma}^2$ методом максимального правдоподобия с точностью до двух знаков после десятичной точки.

42. В модели парной регрессии $R^2 = 0.9$, $TSS = 100$ и 12 наблюдений. Несмещённая оценка дисперсии случайной ошибки равна

a) 1.1
b) 1
c) 0.8
d) 1.3
e) 1.2
f) 0.9

43. Найдите коэффициент при y_{t-2} в выражении $(7 + 7L^2 + (1+L)^3)y_t$.

44. Арамис построил регрессию по 66 наблюдениям:

$$\hat{Y}_i = \underset{(0.4)}{4} + \underset{(5)}{6}X_i + \underset{(2)}{4.4}Z_i - \underset{(2)}{3}Q_i - \underset{(3)}{9}R_i + \underset{(10)}{16}S_i.$$

Показатель R^2_{adj} *может* увеличиться при удалении из модели группы факторов

a) X, Q, S
b) X, S
c) Q, S
d) S
e) X, Q

45. Процесс u_t является белым шумом. Нестационарным является процесс

a) $Y_t = -Y_{t-1} + u_t$
b) $Y_t = 7 + u_t + 0.2u_{t-1} - 1.2u_t - 2$
c) Y_t независимо и одинаково распределены $\mathcal{N}(7;16)$
d) $Y_t = u_t + 2u_{t-1}$
e) $Y_t = 5 + 0.1Y_{t-1} + u_t + 0.2u_{t-1}$

46. Модель $y_i = \beta_1 + \beta_x x_i + u_i$ оценивают с помощью МНК. Известно, что $\bar{x} = 2$, $\bar{y} = 4$, и $\hat{\beta}_x = 5$. Чему равна ордината пересечения линии регрессии с вертикальной осью Oy?

47. У меня набор данных из трёх наблюдений: $x = (0, 1, 1)$, $y = (1, 2, 8)$. Чему равен МНК-прогноз $\hat{y}_3$ в модели $y_i = \beta_1 + \beta x x_i + u_i$?

48. Если $\mathbb{E}(X) = 4$, $\mathbb{E}(Y) = 3$, $\text{Var}(X) = 6$, $\text{Var}(Y) = 7$, $\text{Cov}(X, Y) = -1$, то $\text{Cov}(1 - X + 2Y, X)$ равна

   a) -4
   b) -9
   c) 4
   d) 8
   e) -8
   f) 9

49. Выборочная дисперсия остаточной компоненты ряда равна 100, выборочная дисперсия суммы сезонной и остаточной компоненты равна 150. Найдите силу выраженности сезонности с точностью до двух знаков после десятичной точки.

50. Известно два значения автоковариационной функции стационарного процесса $\gamma_0 = 500$ и $\gamma_7 = 200$. Найдите точное значение автокорреляционной функции ρ_7. Ответ вводите с точностью до двух знаков после десятичной точки.

51. В множественной регрессии с константой матрица X имеет размер 55×5, $TSS = 600$, $ESS = 100$. Несмещённая оценка для дисперсии случайных ошибок модели равна

   a) 30
   b) 5.48
   c) 20
   d) 4.47
   e) 10
   f) 9

52. Укажите число параметров в $ARIMA(2, 2, 2)$ модели с ненулевым математическим ожиданием и нормально распределенными ошибками.

53. Для набора панельных данных истинна спецификация модели со случайными эффектами, однако Вовочка оценивает модель с фиксированными эффектами. Вовочкины оценки коэффициентов β окажутся

   a) несмещёнными и эффективными
   b) состоятельными и неэффективными
   c) смещёнными и неэффективными
   d) состоятельными и эффективными
   e) несостоятельными

54. Используя метод наименьших квадратов в задаче парной регрессии, исследователь

   a) может найти оценки коэффициентов, но не истинные значения коэффициентов.
   b) может найти и оценки коэффициентов, и истинные значения коэффициентов.
   c) может найти истинные значения коэффициентов и построить доверительный интервал для них.
   d) может найти истинные значения коэффициентов, но не оценки коэффициентов.
   e) не может найти оценки коэффициентов, но может построить доверительный интервал для них.

55. По набору данных cars из R оцените модель

$$dist_i = \beta_1 + \beta_2 speed_i + u_i.$$

Ошибки модели нормально распределены и удовлетворяют классическим предпосылкам.
Для машины со стартовой скоростью 20 миль в час постройте 95%-й предиктивный интервал для фактической длины тормозного пути в футах.
В ответе укажите правую границу интервала с точностью до двух знаков после десятичной точки.
Набор данных встроен в R и доступен по ссылке: https://github.com/vincentarelbundock/Rdatasets/raw/maste

56. Если гипотеза $\beta_2 + \beta_3 = 1$ верна, то модель $\ln Y_i = \beta_1 + \beta_2 \ln X_i + \beta_3 \ln Z_i + u_i$ совпадает с моделью

a) $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(X_i/Z_i) + u_i$
b) $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(Y_i/X_i) + u_i$
c) $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(Y_i/Z_i) + u_i$
d) $\ln Y_i = \beta_1 + \beta_2 \ln(X_i/Z_i) + u_i$
e) $\ln Y_i = \beta_1 + \beta_2 \ln(Z_i/Y_i) + u_i$

57. У меня набор данных из трёх наблюдений: $x = (1, 2, 3)$, $y = (5, 6, 10)$.
Чему равна МНК-оценка $\hat{\beta}$ в модели $y_i = 3 + \beta x_i + u_i$?
Ответ вводите с точностью до двух знаков после запятой.

58. Сколько свободных параметров оценивается в $ETS(ANA)$ модели по квартальным данным?

59. Оценка $\hat{\beta}_{2SLS}$ модели $Y = X\beta + u$ получена двухшаговым МНК с матрицей инструментальных переменных Z. Если число инструментов превышает количество включенных в модель факторов, то $\hat{\beta}_{2SLS}$ имеет вид

a) $(Z'X)^{-1}Z'Y$
b) $(X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y$
c) $Z(Z'Z)^{-1}Z'X$
d) $(X'Z)^{-1}Z'Y$
e) $(X'Z(Z'Z)^{-1}Z'X)^{-1}Z'Z(Z'Z)^{-1}$

60. Сколько свободных параметров оценивается в ETS(MAdM) модели на квартальных данных?

61. Основная гипотеза модели адаптивных ожиданий состоит в том, что

a) $X^e_{t+1} - X^e_t = \lambda(X_t - X^e_t),\ 0 \le \lambda < 1$
b) $Y_t - Y_{t-1} = \delta(Y^e_t - Y_{t-1}),\ 0 < \delta \le 1$
c) $X^e_t - X^e_{t-1} = (1-\lambda)(X_t - X_{t-1}),\ 0 \le \lambda < 1$
d) $Y_t - Y^e_{t-1} = \delta(Y^e_t - Y_{t-1}),\ 0 < \delta \le 1$
e) $Y^e_t = \left(1 - \frac{1}{\delta}\right) Y_{t-1} + \frac{1}{\delta} Y_t,\ 0 < \delta \le 1$

62. Винни-Пух пытается понять, от каких переменных может зависеть его потребление мёда. Собрав 100 разных переменных, он построил 100 парных регрессий и проверил в них значимость коэффициента при каждой из переменных на уровне значимости 0.05. Пятачок понимает, что все 100 переменных не имеют никакого отношения к потреблению мёда и на самом деле просто случайные числа. Помогите Пятачку определить, сколько значимых переменных скорее всего найдёт Винни-Пух.

a) Не хватает данных для ответа
b) 5
c) 10
d) 100
e) 0

63. Известно, что $y_t = 5 + u_t + 9u_{t-1}$, где (u_t) — белый шум с дисперсией 7.
Найдите автокорреляцию первого порядка ρ_1 процесса (y_t) с точностью до двух знаков после десятичной точки.

64. Уоррен Баффет проверяет гипотезу H_0: $g(\beta) = 0$ для модели $Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + u_i$ с помощью теста множителей Лагранжа. Для теста Уоррену необходимо знать оценки параметров

   a) регрессии на константу
   b) регрессии на все факторы кроме константы
   c) только модели без ограничений
   d) только модели с ограничениями
   e) модели с ограничениями, и модели без ограничений

65. Рассмотрим два процесса.
   Первый, $y_t = 5 + u_t + 9u_{t-1}$, где (u_t) — белый шум с дисперсией 3.
   Второй, $x_t = 5 + v_t + kv_{t-1}$, где $k \neq 9$ и (v_t) — белый шум с дисперсией σ^2.
   Известно, что у процессов (y_t) и (x_t) одинаковые автоковариационные функции. Найдите k с точностью до двух знаков после десятичной точки.

66. Известно, что $y_1 = 6$, $y_{10} = 30$, а остальные наблюдения — пропущенные.
   С помощью линейной интерполяции восстановите y_2 с точностью до двух знаков после десятичной точки.

67. Рассмотрим $MA(2)$ процесс $y_t = 6 + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2}$, где (u_t) — белый шум. Пусть λ_1 и λ_2 — корни характеристического уравнения MA-части, $\lambda_2 + \alpha_1 \lambda + \alpha_2 = 0$.
   В каком из предложенных случаев процесс (y_t) является обратимым относительно белого шума (u_t)?

   a) $\lambda_1 = 1.3$, $\lambda_2 = 1.8$
   b) $\lambda_1 = 0.3$, $\lambda_2 = -0.8$
   c) $\lambda_1 = 0.3$, $\lambda_2 = 1.8$
   d) нет верного ответа
   e) $\lambda_1 = 1$, $\lambda_2 = -1$
   f) $\lambda_1 = -2$, $\lambda_2 = -7$

68. Рассмотрим лаговые многочлены $P(L) = 1 - 0.2L$ и $Q(L) = 1 + 4L + \alpha L^2$.
   При каком значении α многочлены будут сократимыми?
   Ответ вводите с точностью до двух знаков после десятичной точки.

69. Рассмотрим модель $Y_i = \beta_0 + \beta_z Z_i + \beta_w W_i + u$ при гетероскедастичности. Василий рассчитывает стандартную ошибку $se(\hat{\beta}_w)$ по формуле $se(\hat{\beta}_w) = \sqrt{RSS \cdot (X'X)^{-1}_{33}/(n-3)}$, где X — матрица всех регрессоров.
   Эта стандартная ошибка является

   a) смещённой
   b) смещённой вниз
   c) несмещённой
   d) смещённой вверх
   e) состоятельной

70. Для регрессии $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u$, оцененной по 36 наблюдениям с $R^2 = 0.9$, значение тестовой статистики для проверки гипотезы об адекватности регрессии равно

   a) 32/27
   b) 96
   c) 11/9
   d) невозможно вычислить по имеющимся данным
   e) 99

71. При оценивании регрессионной модели $Y_t = a_0 + \sum_{j=1}^{3} a_j X_{jt} + u_t$ по 20 наблюдениям получено значение статистики Дарбина-Уотсона $d = 0.8$. При уровне значимости 1% это свидетельствует о

   a) Тест Дарбина-Уотсона вообще не проверяет наличие автокорреляции
   b) Отсутствии автокорреляции
   c) 3 и 4
   d) Отрицательной автокорреляции
   e) Попадании в зону неопределенности
   f) Положительной автокорреляции

72. Для модели $Y_i = \beta_1 + \beta_2 X_i + u_i$, где $u_i \sim \mathcal{N}(0, \sigma_u^2)$ тестовая статистика $\frac{\hat{\beta}_2 - \beta_2^0}{se(\hat{\beta}_2)}$ имеет распределение

   a) χ^2_{n-2}
   b) $\mathcal{N}(0,1)$
   c) χ^2_1
   d) t_{n-2}
   e) $\mathcal{N}(0, \sigma_u^2)$

73. Логарифм наблюдаемой величины описывается ETS(AAN) моделью, $\ln y_t \sim ETS(AAN)$. Условная медиана y_{T+1} в два раза меньше условного ожидания.
   Найдите ширину 95%-го предиктивного интервал для $\ln y_{T+1}$ с точностью до двух знаков после запятой.

74. Общеизвестно, что потребление мёда Винни-Пухом зависит, при этом положительно, от количества стихов, сочинённых им за день. К сожалению, Винни-Пух забывчив и всегда называет число сочинённых им стихов с ошибкой. Тогда оценка β_1 в регрессии $Honey_i = \beta_0 + \beta_1 Poems_i + u_i$ окажется

   a) Несостоятельной
   b) Несостоятельной, заниженной
   c) Смещённой, но состоятельной
   d) Несмещенной, но не состоятельной
   e) Несостоятельной, завышенной

75. В парной регрессии величина $\bar{Y} - \hat{\beta}_1 - \hat{\beta}_2 \bar{X}$

   a) не существует
   b) равна 0
   c) может принимать любое неотрицательное значение
   d) равна 1
   e) равна (-1)
   f) может принимать любое положительное значение

76. Рассмотрим модель без константы $y_i = \beta x_i + u_i$, оцениваемую с помощью МНК. Вектор остатков обозначим $\hat{u}$, а вектор из единичек обозначим s.
   Выберите верное утверждение

   a) $\hat{u} \perp s$
   b) $\hat{x} \perp \hat{u}$
   c) $\hat{y} \perp x$
   d) $x \perp u$
   e) $\hat{u} \perp x$
   f) $\hat{y} \perp u$

77. Исследовательница Надежда по пяти наблюдениям оценила регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_x x_i + \hat{\beta}_w w_i$ с помощью МНК.
   И получила вектор остатков $(5, 4, 3, 2, ?)$, а последний остаток Надежда забыла записать.
   Чему равен пропущенный остаток?

78. По данным 570 индивидуумов оценили зависимость почасовой оплаты в долларах, $EARN$, от длительности обучения индивидуума в годах, S, от результата тестирования индивида в баллах, $ASVABC$, и пола индивидуума, $MALE$, (1 — для мужчин, 0 — для женщин):

$$\ln \widehat{EARN}_i = \underset{(0.124)}{0.904} + \underset{(0.056)}{0.01}\, S_i + \underset{(0.002)}{0.0157} ASVABC_i + \underset{(0.1)}{0.27} MALE_i$$

   Почасовая оплата труда мужчин

   a) больше на 0.27 доллара
   b) больше на 0.27 процента
   c) не отличается от оплаты труда женщин
   d) больше на 27 долларов
   e) больше на 27 процентов

79. Найдите коэффициент при y_{t-2} в выражении $(7 + 6L^2 + (1 + L)^4) y_t$.

80. Капитолина оценила регрессию $\hat{y}_i = 0.7 + 2x_i + 3w_i$ с помощью МНК по 10 наблюдениям. Классические стандартные ошибки коэффициентов равны 0.1, 0.2 и 0.3 соответственно.
Наблюдения являются независимыми, на ошибки выполнены классические предпосылки и предпосылка о нормальности, $u_i \sim \mathcal{N}(0; \sigma^2)$, $\mathrm{Cov}(u_i, u_j \mid X) = 0$ для $i \neq j$.
Найдите 80% доверительный интервал для коэффициента β_x.
Можно использовать функции из статистических пакетов или таблицы.
В ответе укажите левую границу интервала с точнотью до двух знаков после десятичной точки.

81. Рассмотрим модель
$$y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i.$$
Все предпосылки классической регрессионной модели выполнены. По 25 наблюдениям оказалось, что сумма квадратов остатков равна $RSS = 60$.
В регрессии z на остальные предикторы оказалось, что $RSS_z = 30$.
Найдите оценку дисперсии $\widehat{\mathrm{Var}}(\hat{\beta}_3 \mid X)$.
Ответ вводите с точность до двух знаков после десятичной точки.

82. По данным для 27 фирм была оценена зависимость выпуска Y от труда L и капитала K с помощью моделей:
$\ln Y_i = b_1 + b_2 \ln L_i + b_3 \ln K_i + u_i$ (1)
$\ln Y_i = b_1 + b_2(\ln L_i + \ln K_i) + u_i$ (2)
Суммы квадратов остатков в этих моделях известны, $RSS_1 = 8$ и $RSS_2 = 10$. F-статистика для проверки гипотезы о равенстве эластичностей по труду и по капиталу равна

a) 6
b) 12
c) 4
d) 2
e) 8

83. В модели множественной регрессии с константой оказалось, что сумма квадратов остатков равна 100, а общая сумма квадратов равна 1000.
Найдите $\sum_i (y_i - \hat{y}_i)^2$.

84. Рассмотрим уравнение
$$y_t = 3 + 0.7 y_{t-1} + u_t + 6u_{t-1},$$
где (u_t) — белый шум.
Определите, являются ли верными утверждения A и B.
A: уравнение имеет одно стационарное решение вида $MA(\infty)$ относительно (u_t).
B: для MA-части уравнения выполнено условие обратимости.

a) A верно, B неверно.
b) A неверно, B верно.
c) A верно, B верно.
d) A неверно, B неверно.

85. Рассмотрим модель
$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + u_i \cdot x_i.$$
Буквой X обозначена матрица всех предикторов. Выполнены классические предпосылки на ошибку $\mathbb{E}(u_i \mid X) = 0$, $\mathrm{Var}(u_i \mid X) = \sigma^2$, $\mathrm{Cov}(u_i \mid X) = 0$.
Найдите условную дисперсию $\mathrm{Var}(y_i \mid X)$.

a) σ^2
b) σx_i
c) $\sigma^2 \cdot x_i^2$
d) x_i
e) u_i^2
f) x_i^2

86. Инструмент Z_t для состоятельной оценки динамической модели $Y_t = \alpha + \beta_0 X_t + \beta_1 Y_{t-1} + u_t$, где $u_t = u_t + \lambda_1 u_{t-1} + \lambda_2 u_{t-2}$,

    a) удовлетворяет условию $\text{Corr}(Z_t, Y_{t-1}) = 0$
    b) удовлетворяет условию $\text{Corr}(Z_t, X_t) = 0$
    c) удовлетворяет условию $\text{Corr}(Z_t, u_t) \to 1$
    d) удовлетворяет условию $\text{Corr}(Z_t, Y_{t-1}) \neq 0$
    e) не требуется

87. Рассмотрим алгоритм градиентного бустинга над решающими деревьями для задачи регрессии. Выберите верное утверждение о том, какие наблюдения используются для обучения каждого дерева.

    a) нет верного ответа
    b) каждое дерево использует свою бутстрэп-реплику исходных наблюдений
    c) наблюдения распределяются примерно поровну между деревьями случайным образом
    d) каждое дерево использует все исходные наблюдения
    e) каждое последующее дерево использует всё больше и больше наблюдений
    f) каждое последующее дерево использует всё меньше и меньше наблюдений

88. Перед Винни-Пухом временной ряд потребления мёда без тренда и сезонными колебаниями постоянной амплитуды. Какая модель из предложенных лучше подходит для описания данного ряда?

    a) ETS(MNM)
    b) ETS(AAN)
    c) ETS(MAM)
    d) ETS(AAA)
    e) ETS(ANA)
    f) ETS(AAdA)

89. Какая модель простому экспоненциальному сглаживанию?

    a) ETS(ANA)
    b) ETS(AAdA)
    c) ETS(ANN)
    d) ETS(AAN)
    e) ETS(AAA)
    f) нет верного ответа

90. По $n = 450$ 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 $\hat{\alpha}_2 = 20$ и $se(\hat{\alpha}_2) = 5$. Согласно этим данным, на уровне значимости 5% гипотеза о

    a) пропущенной переменной $1/Z_i$ отвергается
    b) гомоскедастичности отвергается
    c) пропущенной переменной $1/Z_i$ не отвергается
    d) верной функциональной форме отвергается
    e) гомоскедастичности не отвергается
    f) верной функциональной форме не отвергается

91. Выберите верное утверждение про модели бинарного выбора.

a) Модели бинарного выбора предназначены для выбора одного из двух возможных предикторов.
b) Модели бинарного выбора предназначены для включения предиктора, принимающего два значения.
c) Модели бинарного выбора предназначены для моделирования зависимой переменной, принимающей счетноё число значений.
d) Модели бинарного выбора предназначены для выбора одной из двух возможных линейных моделей.
e) Модели бинарного выбора предназначены для моделирования зависимой переменной, принимающей два значения.
f) Модели бинарного выбора предназначены для моделирования зависимой переменной с пропущенными значениями.

92. Если процесс является стационарным в широком смысле, то

a) он является стационарным в узком смысле
b) для него выполняется основная гипотеза в тесте Дикки-Фуллера
c) его приращения являются стационарным процессом
d) его автоковариационная функция является постоянной
e) это AR процесс с корнями характеристического уравнения, меньшими 1
f) это белый шум

93. Если нулевая гипотеза в расширенном тесте Дики-Фуллера с константой отвергается, то исходный ряд можно считать

a) стационарным в первых разностях
b) стационарным во вторых разностях
c) нестационарным
d) возрастающим
e) стационарным

94. В исходной выборке 9 наблюдений. Найдите вероятность того, что второе наблюдение из исходной выборки попадёт в очередную бутстрэп выборку ровно 3 раз.
Ответ укажите с точностью до двух знаков после десятичной точки.

95. У Маши две монетки: медная и серебряная. Маша подкинула каждую монетку 100 раз. Затем с помощью метода максимального правдоподобия Маша трижды оценила вероятность выпадения орла: для медной монетки, для серебряной и по объединённой выборке. Значения функции правдоподобия равны $\ell_{copper} = -300$, $\ell_{silver} = -200$ и $\ell_{common} = -510$.
LR статистика, проверяющая гипотезу о равенстве вероятностей выпадения орла для двух монеток, равна

a) 10
b) 1010
c) 5
d) 500
e) 20

96. Исследователь Феофан оценил с помощью МНК модель $Y = \beta_0 I + \beta_1 Z + \beta_2 W + u$, где I — столбец из единиц. Для матрицы факторов, $X = (IZW)$, известно, что

$$(X'X)^{-1} = \begin{pmatrix} 0.04 & 0.012 & -0.008 \\ 0.012 & 0.03 & -0.007 \\ -0.008 & -0.007 & 0.02 \end{pmatrix}$$

Предпосылки теоремы Гаусса-Маркова выполнены. Отношение дисперсии оценки $\hat{\beta}_0$ к дисперсии оценки $\hat{\beta}_2$ равно

a) $3/2$
b) 2
c) $1/2$
d) $-5/1$
e) $10/3$

97. Рассмотрим модель, стоящую за тета-методом,

$$\begin{cases} y_t = \ell_{t-1} + 2 + u_t \\ \ell_t = \ell_{t-1} + 2 + 0.3 u_t \end{cases}.$$

Перепишите эту модель в виде

$$\Delta y_t = \theta_1 + \theta_2 u_{t-1} + u_t.$$

В ответ укажите коэффициент θ_2.

98. Исследователь Винни-Пух оценил регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$ с помощью LASSO с очень-очень большим штрафом.
Известно, что $\bar{x} = 10$, $\bar{y} = 20$, $n = 100$.
Найдите полученную им оценку $\hat{\beta}_2$.

99. Какие последствия имеет нестрогая мультиколлинеарность?

a) Оценки коэффициентов остаются несмещёнными, однако доверительные интервалы шире по сравнению с отсутствием мультиколлинеарности.
b) Оценки коэффициентов становятся несостоятельными.
c) Предиктивные интервалы резко сужаются по сравнению с отсутствием мультиколлинеарности.
d) Предиктивные интервалы резко расширяются по сравнению с отсутствием мультиколлинеарности.
e) Оценки коэффициентов становятся смещёнными.
f) Оценки коэффициентов остаются несмещёнными, однако доверительные интервалы уже по сравнению с отсутствием мультиколлинеарности.

100. Василий сменил единицы измерения ряда с рублей на тысячи рублей.
На какое число при этом домножится средняя абсолютная процентная ошибка наивной модели?

101. 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) $(I'I)^{-1} I'Y$
b) $\bar{Y}$
c) $\frac{\sum Y_i/X_i}{\sum 1/X_i}$
d) $\frac{\sum Y_i X_i}{\sum X_i^2}$
e) $\frac{\sum Y_i/X_i}{\sum 1/X_i^2}$
f) $\frac{\sum Y_i X_i}{\sum X_i}$

102. Выберите верное утверждение про мультиколлинеарность.

   a) Мультиколлинеарность — это линейная зависимость между регрессорами.
   b) Мультиколлинеарность — это нелинейная зависимость между регрессорами.
   c) Мультиколлинеарность — это линейная зависимость между зависимой переменной и случайной ошибкой.
   d) Мультиколлинеарность — это линейная зависимость между регрессорами и случайной ошибкой.
   e) Мультиколлинеарность — это нелинейная зависимость между зависимой переменной и случайной ошибкой.
   f) Мультиколлинеарность — это линейная зависимость между зависимой переменной и регрессорами.

103. Тест Саргана для проверки валидности инструментов можно использовать только в том случае, если число инструментов

   a) больше числа эндогенных переменных
   b) совпадает с числом эндогенных переменных
   c) совпадает с числом экзогенных переменных
   d) меньше числа экзогенных переменных
   e) меньше числа эндогенных переменных

104. Капитолина оценила регрессию $\hat{y}_i = 0.7 + 2x_i + 3w_i$ с помощью МНК по 1000 наблюдениям. Она хочет проверить гипотезу $\beta_x = \beta_w$, однако не хочет ничего рассчитывать руками. Как поступить Капитолине, чтобы легко проверить желаемую гипотезу по стандартной табличке, выдаваемой статистическими программами?

   a) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i - w_i) + \hat{\beta}_3 w_i$, посмотреть на значимость коэффициента $\hat{\beta}_2$.
   b) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i + w_i) + \hat{\beta}_3 w_i$, посмотреть на значимость коэффициента $\hat{\beta}_2$.
   c) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i + w_i) + \hat{\beta}_3 w_i$, посмотреть на значимость коэффициента $\hat{\beta}_3$.
   d) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i + w_i) + \hat{\beta}_3(x_i - w_i)$, посмотреть на значимость коэффициента $\hat{\beta}_2$.
   e) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i - w_i) + \hat{\beta}_3(w_i - x_i)$, посмотреть на значимость коэффициента $\hat{\beta}_3$.
   f) Оценить регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2(x_i - w_i) + \hat{\beta}_3 w_i$, посмотреть на значимость коэффициента $\hat{\beta}_3$.

105. Выберите верное утверждение про оценки коэффициентов логит-модели.

   a) Оценки коэффициентов не являются случайными.
   b) Оценки коэффициентов являются асимптотически нормальными.
   c) Доверительный интервал для коэффициентов строится с помощью t-распределения.
   d) Оценки коэффициентов строго больше соответствующих оценок модели линейной регрессии.
   e) Оценки коэффициентов строго меньше соответствующих оценок модели линейной регрессии.

106. На панельных данных коэффициенты при факторах, постоянных во времени, НЕ могут быть оценены с помощью

a) метода максимального правдоподобия
b) RE-оценки
c) МНК для модели с полным набором дамми-переменных для каждого индивида
d) сквозной регрессии
e) between-регрессии

107. Капитолина оценила регрессию $\hat{y}_i = 0.7 + 2x_i + 3w_i$ с помощью МНК по 1000 наблюдениям. Оценённая ковариационная матрица оценок коэффициентов равна

$$\mathrm{Var}(\hat{\beta}) = \begin{pmatrix} 4 & 1 & -1 \\ 1 & 9 & 0 \\ -1 & 0 & 16 \end{pmatrix}.$$

Наблюдения являются независимыми, на ошибки выполнены классические предпосылки, $\mathbb{E}(u_i \mid X) = 0$, $\mathrm{Var}(u_i \mid X) = \sigma^2$, $\mathrm{Cov}(u_i, u_j \mid X) = 0$ для $i \neq j$.
Постройте 95%-й доверительный интервал для разницы $\beta_x - \beta_w$.
Можно использовать функции из статистических пакетов или таблицы.
В ответ введите правую границу с точностью до двух знаков после десятичной точки.

108. При работе с панельными данными для выбора между моделью с фиксированными эффектами и моделью со случайными эффектами используется

a) тест отношения правдоподобия
b) поиск на сетке
c) тест Голдфелда-Квандта
d) тест Хаусмана
e) тест Бройша–Пагана

109. У Агнессы 120 наблюдений и она оценивает качество прогнозов на один шаг вперёд с помощью кросс-валидации скользящим окном. Ширина окна равна 70.
Сколько раз будет оценена каждая модель при выполнении кросс-валидации?

110. Пантелеймон оценил парную регрессию по 50 наблюдениям

$$\hat{y}_i = 8.9 + 12.9x_i.$$

Оценка ковариационной матрицы оценок коэффициентов имеет вид

$$\begin{pmatrix} 16.7 & -4.6 \\ -4.6 & 1.74 \end{pmatrix}.$$

Величина суммы квадратов остатков равна 10871.
Предпосылки классической линейной модели выполнены, ошибки имеют нормальное распределение.
Оцените дисперсию прогноза для $x = 1$.
Ответ вводите с точностью до двух знаков после десятичной точки.

111. Пантелеймон оценил парную регрессию по 50 наблюдениям

$$\hat{y}_i = 8.9 + 12.9x_i.$$

Оценка дисперсии прогноза при $x = 1$ равна 100. Оценка дисперсии случайной ошибки равна 400. Предпосылки классической линейной модели выполнены, ошибки имеют нормальное распределение.
Постройте 95%-й предиктивный интервал для y_i при $x_i = 1$.
В ответе укажите правую границу интервала с точностью до двух знаков после десятичной точки.

112. При проверке гипотезы о значимости коэффициента линейной регрессии p-значение, соответствующее тестовой статистике, оказалось равным 0.07. Отсюда следует, что

a) длина 95% доверительного интервала для этого коэффициента больше 0.07
b) длина 95% доверительного интервала для этого коэффициента меньше 0.07
c) соответствующий коэффициент не значим при уровне значимости 5%
d) соответствующий коэффициент значим при уровне значимости 1%
e) длина 95% доверительного интервала для этого коэффициента равна 0.07

113. У меня набор данных из трёх наблюдений: $x = (2, 2, 2)$, $y = (3, 6, 12)$.
Чему равен МНК-прогноз $\hat{y}_3$ в модели $y_i = \beta x_i + u_i$?

114. Сколько свободных параметров оценивается в ETS(MNM) модели на квартальных данных?

115. У меня набор данных из трёх наблюдений: $x = (1, 2, 3)$, $y = (5, 6, 10)$.
Чему равна МНК-оценка $\hat{\beta}$ в модели $y_i = \beta x_i + u_i$?
Ответ вводите с точностью до двух знаков после запятой.

116. В рамках ETS(ANN) модели найдите точечный прогноз на три шага вперёд, Последнее значение уровня (сглаженного ряда) равно 200, $\ell_t = \ell_{t-1} + 0.3u_t$, $u_t \sim \mathcal{N}(0; 16)$, $y_t = \ell_{t-1} + u_t$.
Ответ введите с точностью до двух знаков после десятичной точки.

117. Исследовательница Клеопатра оценила модель $\ln Y_i = \beta_0 + \beta_1 \ln X_i + \beta_2 \ln Z_i + \beta_3 \ln W_i + u_i$. Клеопатра хочет протестировать гипотезу H_0: $\beta_3 + 2\beta_1 = 1$. Для этой цели можно оценить вспомогательную регрессию

a) $\ln(Y_i/W_i) = \gamma_0 + \gamma_1 \ln(X_i/W_i^2) + \gamma_2 \ln Z_i + u_i$
b) $\ln(Y_i/W_i) = \gamma_0 + \gamma_1 \ln(X_i \cdot W_i^2) + \gamma_2 \ln Z_i + u_i$
c) $\ln(Y_i/W_i^2) = \gamma_0 + \gamma_1 \ln(X_i/W_i) + \gamma_2 \ln Z_i + u_i$
d) $\ln(Y_i \cdot W_i) = \gamma_0 + \gamma_1 \ln(X_i/W_i^2) + \gamma_2 \ln Z_i + u_i$
e) $\ln(Y_i \cdot W_i) = \gamma_0 + \gamma_1 \ln(X_i \cdot W_i^2) + \gamma_2 \ln Z_i + u_i$

118. Рассмотрим две конкурирующие модели: $ARMA(2, 1)$ с ненулевым ожиданием и $ARIMA(0, 1, 2)$ с нулевым ожиданием.
Каким способом разумно выбрать наилучшую из них?

a) нет верного ответа
b) ADF тест без константы
c) KPSS тест с константой
d) KPSS тест с трендом
e) ADF тест с трендом
f) критерий Акаике

119. Для регрессионной модели со свободным членом известно, что

$$X'X = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 3 & 5 \end{pmatrix}, \quad X'Y = \begin{pmatrix} 40 \\ 10 \\ 13 \end{pmatrix}, \quad \sum_{i=1}^{n} Y_i^2 = 140.$$

Коэффициент R^2 в этой модели равен

a) недостаточно информации, чтобы вычислить R^2
b) 0.5
c) $9/35$
d) 0.6
e) $13/14$

120. Рассеянная исследовательница Надежда ошибочно оценила модель $y_i = \beta_1 + \beta_x x_i + \beta_w w_i + u_i$ и получила неверные оценки $\hat{\beta}$.
А затем, используя неверные оценки коэффициентов, посчитала $\sum \hat{u}_i^2$ и $\sum \hat{u}_i$.
Сравните эти две суммы с верно рассчитанными.

a) $\sum \hat{u}_i^2$ занижена, а $\sum \hat{u}_i$ завышена.
b) $\sum \hat{u}_i^2$ может быть и завышена, и занижена, а $\sum \hat{u}_i$ занижена.
c) обе суммы могут быть и завышены, и занижены.
d) $\sum \hat{u}_i^2$ завышена, а $\sum \hat{u}_i$ может быть и завышена, и занижена.
e) $\sum \hat{u}_i^2$ завышена, а $\sum \hat{u}_i$ занижена.
f) $\sum \hat{u}_i^2$ может быть и завышена, и занижена, а $\sum \hat{u}_i$ завышена.

121. У исследователя всего четыре наблюдения, $y = (1, 3, 5, 7)$.
Найдите общую сумму квадратов TSS.

122. Выберите верное утверждение про главные компоненты.

a) Главные компоненты сохраняют среднее значение исходных переменных.
b) Первая главная компонента совпадает с зависимой переменной.
c) Первая главная компонента имеет наименьшую выборочную дисперсию.
d) Главные компоненты некоррелированны.
e) Главные компоненты имеют единичную длину.
f) Первая главная компонента имеет наибольшее среднее значение.

123. Пантелеймон включил в модель 200 совершенно бессмысленных регрессоров, которые никак не влияют на переменную y. Пантелеймон не знает этого и проверяет значимость коэффициента при каждом коэффициенте на уровне значимости $\alpha = 0.02$.
Сколько ложно значимых регрессоров в среднем обнаружит Пантелеймон?

124. Для регрессии $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 Z_i + \hat{\beta}_3 W_i$, оценённой по 24 наблюдениям, $R^2 = 0.9$.
При проверке гипотезы о неадекватности модели F-статистика равна

a) $189/2$
b) 60
c) $5/9$
d) 45
e) 200.27

125. В парной регрессии на уровне значимости 5%-ов гипотеза H_0: $\beta_2 = 2016$ не отвергается. Из этого можно сделать вывод, что на соответствующем уровне значимости

a) H_a: $\beta_2 \neq 0$ отвергается
b) нет верного ответа
c) H_a: $\beta_2 \neq 0$ не отвергается
d) доверительный интервал для β_2 не содержит ноль
e) H_0: $\beta_2 = 0$ отвергается
f) H_0: $\beta_2 = 0$ не отвергается

126. Рассмотрим уравнение

$$y_t = 3 + 0.7y_{t-1} + u_t + 0.6u_{t-1},$$

где (u_t) — белый шум.
Определите, являются ли верными утверждения A и B.
A: уравнение имеет одно стационарное решение вида $MA(\infty)$ относительно (u_t).
B: для MA-части уравнения выполнено условие обратимости.

a) A верно, B верно.
b) A верно, B неверно.
c) A неверно, B неверно.
d) A неверно, B верно.

127. Выборочная корреляция между регрессорами X и Z равна 0.5. В регрессии $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 Z_i$ показатель VIF для регрессора X равен

a) $1/2$
b) 2
c) $3/4$
d) $4/3$
e) $1/4$

128. Использование скорректированных стандартных ошибок Уайта при гомоскедастичности приведет к

a) получению состоятельной оценки дисперсии случайной ошибки
b) понижению эффективности МНК оценок коэффициентов
c) смещённости МНК оценок коэффициентов
d) несостоятельности МНК оценок коэффициентов
e) повышению эффективности МНК оценок коэффициентов

129. Рассмотрим модель множественной регрессии $Y = X\beta + u$, где $\hat{Y} = X\hat{\beta}$, $e = Y - \hat{Y}$. Величина RSS — это квадрат длины вектора

a) $\hat{Y} - \bar{Y}$
b) e
c) $\hat{Y}$
d) u
e) $Y - \bar{Y}$

130. Коэффициент R^2 может быть представлен в виде

a) $\sum_{j=2}^{k} \beta_j \frac{\mathrm{sVar}(Y)}{\mathrm{sCov}(X_j, Y)}$
b) $\sum_{j=2}^{k} \hat{\beta}_j \frac{\mathrm{sCov}(X_j, Y)}{\mathrm{sVar}(Y)}$
c) $\sum_{j=2}^{k} \hat{\beta}_j \frac{\mathrm{sCorr}(X_j, Y)}{\mathrm{sVar}(\hat{Y})}$
d) $\sum_{j=2}^{k} \hat{\beta}_j \frac{\mathrm{sCorr}(X_j, Y)}{\mathrm{sVar}(Y)}$
e) $\sum_{j=2}^{k} \beta_j \frac{\mathrm{sCorr}(X_j, Y)}{\mathrm{sVar}(\hat{Y})}$
f) $\sum_{j=2}^{k} \hat{\beta}_j \frac{\mathrm{sVar}(Y)}{\mathrm{sCov}(X_j, Y)}$

131. Рассмотрим модель
$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + u_i.$$
Буквой X обозначена матрица всех предикторов. Выполнены классические предпосылки на ошибку $\mathbb{E}(u_i \mid X) = 0$, $\mathrm{Var}(u_i \mid X) = \sigma^2$, $\mathrm{Cov}(u_i \mid X) = 0$.
Найдите условное ожидание $\mathbb{E}(y_i \mid X)$.

a) $\beta_1 + \beta_2 x_i + \beta_3 x_i^2 + u_i$
b) $\hat{y}_i$
c) $\hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\beta}_3 x_i^2$
d) $\beta_1 + \beta_2 x_i$
e) 0
f) $\beta_1 + \beta_2 x_i + \beta_3 x_i^2$

132. В модели парной регрессии выборочная корреляция между зависимой переменной и предиктором равна -0.7.
Найдите выборочную корреляцию между зависимой переменной и прогнозами.

133. Случайные величины r и s независимы и равновероятно равны 0 или 1. Чему равна условная дисперсия $\mathrm{Var}(r + s^2 \mid r)$?

a) $r + 0.25$
b) r
c) 1
d) $r + 0.5$
e) 0.5
f) 0.25

134. Рассмотрим ETS модель с мультипликативной сезонностью для месячных данных. Известно, что $s_5 = 0.8$, $u_{17} = 1$, $\gamma = 0.3$.
На сколько процентов сезонный эффект увеличивает значение наблюдаемого ряда в момент времени $t = 17$?
В ответе укажите целое число процентов.

135. Если оценивается модель $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$, а истинной является модель $Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + u_i$, то МНК-оценка $\hat{\beta}_2$ оказывается

a) всегда несмещённой
b) несмещённой при $\beta_1 = 0$
c) несмещённой при нулевой выборочной ковариации X и Z
d) эффективной
e) всегда смещённой
f) равной нулю

136. По 20 наблюдениям Чебурашка оценил модель $Y_i = \beta_0 + \beta_1 X_i + u_i$. Известно, что $\sum X_i = -10$, $\sum X_i^2 = 40$, $\sum X_i Y_i = 10$, $\sum Y_i = 50$.
Сумма оценок МНК коэффициентов $\hat{\beta}_0 + \hat{\beta}_1$ равна

a) 3
b) 5
c) 4
d) 1
e) 2

137. Ряд (u_t) — белый шум с дисперсией равной 25.
Найдите дисперсию $\mathrm{Var}(2u_t + 2u_{t-1} + \sqrt{2}u_{t-2})$.

138. Корни характеристического многочлена AR-части уравнения равны $\lambda_1 = 0.02$ и $\lambda_2 = 0.28$.
Найдите наименьший корень лагового многочлена AR-части.
Ответ вводите с точностью до двух знаков после десятичной точки.

139. Рассмотрим два процесса.
Первый, $y_t = 5 + u_t + 2u_{t-1}$, где (u_t) — белый шум с дисперсией 8.
Второй, $x_t = 5 + v_t + kv_{t-1}$, где $k \neq 2$ и (v_t) — белый шум с дисперсией σ^2.
Известно, что у процессов (y_t) и (x_t) одинаковые автоковариационные функции. Найдите σ^2 с точностью до двух знаков после десятичной точки.

140. Рассмотрим логистическую регрессию с пятью регрессорами помимо константы, оцениваемую методом максимального правдоподобия по n наблюдениям. Статистика $\hat{\beta}_3/se(\hat{\beta}_3)$ для проверки значимости коэффициента β_3 имеет

    a) t-распределение с $n-5$ степенями свободы
    b) t-распределение с $n-6$ степенями свободы
    c) t-распределение с n степенями свободы
    d) асимптотически нормальное распределение
    e) χ^2-распределение с одной степенью свободы

141. Процесс u_t является белым шумом. Нестационарным является процесс

    a) $Y_t = -1Y_{t-1} + u_t$
    b) $Y_t = u_t + 0.1u_{t-1} + 1.5u_{t-2}$
    c) $Y_t = 0.1Y_{t-1} + u_t$
    d) $Y_t = 2017$
    e) $Y_t = 2017u_t$

142. Обобщенный МНК служит для оценивания регрессионных моделей в случае нарушений следующего условия теоремы Гаусса-Маркова:

    a) $\text{rank}X = k$
    b) $\text{Var}(u) = \sigma^2 I$
    c) Величина Y_i линейна по $\beta_1, \beta_2, \ldots$
    d) $\mathbb{E}(u_i) = 0$
    e) u_i распределены нормально

143. При выполненных условиях регулярности оценки метода максимального правдоподобия могут **НЕ** являться

    a) состоятельными
    b) асимптотически эффективными
    c) асимптотически нормальными
    d) несмещёнными
    e) инвариантными

144. Выберите верное утверждение о ситуации, в которой оценки логит-модели не существуют.

    a) Нужно рассмотреть пробит-модель вместо логит-модели.
    b) Такая ситуация невозможна, оценки максимального правдоподобия всегда существуют.
    c) Разумно исключить из модели константу.
    d) Разумно включить квадраты и кросс-произведения исходных переменных.
    e) Проблему разумно решить введением в целевую функцию дополнительного штрафа за отклонение оценок коэффициентов от нуля.

145. Исследовательница Надежда оценила регрессию в отклонениях, $\hat{y}_i = x_i + 2z_i$ с помощью МНК. Известно, что $\bar{Y} = 5$, $\bar{X} = 6$, $\bar{Z} = -2$. В регрессии нецентрированных переменных, $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 Z_i$, оценка коэффициента $\hat{\beta}_0$ равна

    a) 2
    b) 3
    c) 5
    d) 1
    e) 4

146. Известны средние $\bar{x} = 3$, $\bar{y} = 2$ и скалярное произведение $\langle x, y\rangle = 9$.
    Определим $a_i = x_i - \bar{x}$ и $b_i = y_i - \bar{y}$.
    Чему равно скалярное произведение $\langle a, b\rangle$?

147. У $ARMA(4,1)$-модели с ненулевым ожиданием и нормально распределенными ошибками логарифм правдоподобия оказался равен -190.
    Найдите значение информационного критерия Акаике.

148. Исследователь Борис оценил параметры нескольких моделей:

Модель	Уравнение
1	$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$
2	$\ln Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$
3	$Y = \beta_1 + \beta_4 X_4 + \beta_5 X_5 + u$
4	$Y/X_2 = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$

С помощью R^2_{adj} можно выбрать лучшую из пар моделей

a) 2 и 4
b) 1 и 4
c) 1 и 3
d) 2 и 3
e) 1 и 2
f) 3 и 4

149. Рассмотрим ETS модель с мультипликативной сезонностью и мультипликативной ошибкой для месячных данных. Известно, что $s_5 = 0.8$, $u_{17} = 1.2$, $u_{29} = -1$, $\gamma = 0.5$.
На сколько процентов сезонный эффект уменьшает значение наблюдаемого ряда в момент времени $t = 41$?
В ответе укажите целое число процентов.

150. Выберите верное утверждение про стационарный $ARMA(3, 2)$ процесс.

a) нет верного ответа
b) автокорреляционная функция зануляется начиная с лага 3.
c) автокорреляционная функция зануляется начиная с лага 4.
d) частная автокорреляционная функция зануляется начиная с лага 4.
e) частная автокорреляционная функция зануляется начиная с лага 2.
f) частная автокорреляционная функция зануляется начиная с лага 3.

151. Выберите стационарный процесс (уравнение, которое имеет стационарное решение, не заглядывающее в будущее), если u_t — белый шум.

a) $y_t = 1.2y_{t-1} + u_t$
b) $y_t = u_1 + u_2 + \ldots + u_t$
c) $y_t = tu_t$
d) $y_t = 1.2 + 0.8y_{t-1} - 0.12y_{t-2} + u_t - 3u_{t-1}$
e) $y_t = 2 + 3t + u_t + u_{t-1}$

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

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

имеет вид:

a) $\hat{\lambda}_{ML} = \frac{X_1^2+\ldots+X_n^2}{n}$
b) $\hat{\lambda}_{ML} = -\frac{\ln X_1+\ldots+\ln X_n}{n}$
c) $\hat{\lambda}_{ML} = \frac{X_1+\ldots+X_n}{n}$
d) $\hat{\lambda}_{ML} = \frac{\ln X_1+\ldots+\ln X_n}{n}$
e) $\hat{\lambda}_{ML} = -\frac{X_1+\ldots+X_n}{n}$

153. Что показывает коэффициент вздутия дисперсии VIF оценки коэффициента регрессии?

a) Во сколько раз дисперсия оценки коэффициента больше по сравнению с идеальной ситуацией с некоррелированными предикторами.

b) Во сколько раз дисперсия оценки коэффициента больше по сравнению с идеальной ситуацией со стандартизированными предикторами.

c) Во сколько раз дисперсия оценки коэффициента больше по сравнению с идеальной ситуацией с одинаковыми предикторами.

d) Во сколько раз дисперсия оценки коэффициента больше по сравнению с идеальной ситуацией с несмещёнными предикторами.

e) Во сколько раз дисперсия оценки коэффициента больше по сравнению с идеальной ситуацией парной регрессии.

154. Известно значение автоковариационной функции стационарного процесса $\gamma_9 = 120$ и значение его автокорреляционной функции $\rho_9 = 0.6$.
Найдите дисперсию случайного процесса.

155. Известно, что $y_3 = 2$, а соответствующий остаток равен 5.
Чему равен прогноз $\hat{y}_3$ в модели множественной регрессии?

156. Какая зависимость математического ожидания исходного процесса от времени предполагается в нулевой гипотезе ADF-теста с константой?

a) $\mathbb{E}(y_t)$ строго возрастает

b) $\mathbb{E}(y_t) = \mu + \alpha t$

c) $\mathbb{E}(y_t) = \mu$

d) $\mathbb{E}(y_t) = \mu + \alpha t + \beta t^2$

e) $\mathbb{E}(y_t)$ строго монотонна

f) нет верного ответа

157. Для регрессии $Y_i = \beta_1 + \beta_2 X_i + u_i$, оценённой по 30 наблюдениям с суммой квадратов остатков, равной 15, несмещенная оценка дисперсии случайной составляющей равна

a) $15/32$

b) $15/28$

c) $13/30$

d) $13/28$

e) 2

f) 0.5

158. Случайная величина X имеет t-распределение с пятью степенями свободы.
Найдите вероятность $\mathbb{P}(X > 1)$.
Можно использовать статистические функции в R/Python или других программах. Ответ вводите с точностью до двух знаков после десятичной точки.

159. В регрессии с константой и тремя объясняющими переменными сумма квадратов остатков равна 162, а число наблюдений равно 31. Точечная оценка дисперсии случайной составляющей равна

a) 2.83

b) 7

c) 8

d) 2.45

e) 2.65

f) 6

160. Использование робастных стандартных ошибок в форме Уайта при гетероскедастичности позволяет

a) получить эффективные оценки коэффициентов

b) увеличить точность прогнозов

c) строить корректные доверительные интервалы для коэффициентов

d) устранить смещённость оценок коэффициентов

e) сузить доверительные интервалы для коэффициентов

161. Эконометресса Агнесса была очень недовольна тем, что оценили регрессию

$$\hat{y}_i = 5.42 + 3.25 \cdot x_i - 2.51 male_i.$$

Ведь переменная $male_i$ равна единице для мужчин и нулю для женщин.
Поэтому Агнесса переоценила модель с помощью переменной $female_i$, которая равна нулю для мужчин и единице для женщин:

$$\hat{y}_i = \hat{\alpha}_1 + \hat{\alpha}_2 x_i + \hat{\alpha}_3 female_i.$$

Какую оценку $\hat{\alpha}_1$ получила Агнесса?
Ответ вводите с точностью до двух знаков после десятичной точки.

162. Оценка МНК неизвестного параметра θ для модели $Y_i = \theta X_{1i} + (1+\theta)X_{2i} + u_i, i = 1, \ldots, n$ равна

a) $\frac{\sum_{i=1}^n (X_{1i}+X_{2i})(Y_i-X_{2i})}{\sum_{i=1}^n (X_{1i}+X_{2i})^2}$

b) $\frac{\sum_{i=1}^n (X_{1i}-Y_i)(Y_i-X_{2i})}{\sum_{i=1}^n (X_{1i}+X_{2i})^2}$

c) $\frac{\sum_{i=1}^n (X_{1i}+X_{2i})(Y_i-X_{2i})}{\sum_{i=1}^n (X_{1i}-Y_i)^2}$

d) $\frac{\sum_{i=1}^n (X_{1i}+X_{2i})(Y_i-X_{2i})}{\sum_{i=1}^n (Y_i-X_{2i})^2}$

e) $\frac{\sum_{i=1}^n (X_{1i}+X_{2i})(Y_i-X_{1i})}{\sum_{i=1}^n (X_{1i}+X_{2i})^2}$

163. Метод наименьших квадратов в задаче парной регрессии подбирает оценки коэффициентов так, чтобы

a) прогнозы зависимой переменной были наименее похожими на фактические значения регрессора

b) прогнозы регрессора были наиболее похожими на фактические значения регрессора

c) прогнозы зависимой переменной были наиболее похожими на фактические значения регрессора

d) прогнозы зависимой переменной были наименее похожими на фактические значения зависимой переменной

e) прогнозы зависимой переменной были наиболее похожими на фактические значения зависимой переменной

164. Имеются данные по доходу жены, мужа и продолжительности брака. Доход семьи складывается из дохода жены и мужа. Вася оценил зависимость дохода семьи от продолжительности брака и получил регрессию $\hat{Y}_i = 20 + 3X_i$, Петя оценил зависимость дохода мужа от продолжительности брака и получил регрессию $\hat{Y}_i = 10 + 2X_i$. Маша оценивает зависимость дохода жены от продолжительности брака. Она получит регрессию:

a) $\hat{Y}_i = 20 + 3X_i$

b) недостаточно данных для ответа

c) $\hat{Y}_i = 15 + 2.5X_i$

d) $\hat{Y}_i = 10 + X_i$

e) $\hat{Y}_i = 30 + 5X_i$

f) $\hat{Y}_i = 10 - X_i$

165. При отсутствии автокорреляции в регрессии по n наблюдениям статистика Дарбина-Уотсона имеет

a) нет верного ответа

b) $\mathcal{N}(\mu; \sigma^2)$-распределение

c) t_n-распределение

d) $\mathcal{N}(0; 1)$-распределение

e) $F_{k,n}$-распределение

f) t_{n-k}-распределение

166. Рассмотрим алгоритм случайного леса для задачи регрессии.
Выберите верное утверждение о том, какие наблюдения используются для обучения каждого дерева.

a) каждое дерево использует свою бутстрэп-реплику исходных наблюдений
b) каждое последующее дерево использует всё меньше и меньше наблюдений
c) каждое дерево использует все исходные наблюдения
d) каждое последующее дерево использует всё больше и больше наблюдений
e) наблюдения распределяются примерно поровну между деревьями случайным образом
f) нет верного ответа

167. Леонардо оценил модель логистической регрессии $\hat{\mathbb{P}}(y_i = 1) = \Lambda(2 + 0.05x_i - 0.07z_i)$.
Найдите предел прогноза вероятности $\hat{\mathbb{P}}(y_i = 1)$ при z_i стремящемся к плюс бесконечности.

168. Выберите утверждение про модель локальной линейной регрессии LOESS.

a) В качестве ядерной функции разумно взять периодическую функцию.
b) Существует единственный канонический вариант выбора ядерной функции.
c) LOESS можно использовать только на перекрестных данных.
d) В качестве ядерной функции разумно взять функцию дающую меньший вес для наблюдений вдали от рассматриваемого значения аргумента.
e) Ядерная функция должна принимать как положительные, так и отрицательне значения.
f) Обычная парная регрессия не может быть реализована как частный случай LOESS.

169. Как изменятся оценки логистической регрессии, если заменить у зависимой переменной нули единицами, а единицы — нулями?

a) Оценки коэффициентов увеличатся.
b) Оценки коэффициентов не изменятся.
c) Оценки коэффициентов поменяют знак.
d) Новые оценки коэффициентов будут обратными к старым.
e) Оценки коэффициентов уменьшатся.
f) Оценки коэффициентов изменятся непредсказуемым образом.

170. Сулейман оценил две модели множественной регрессии по 304 наблюдениям. Модель А:

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \beta_4 w_i + u_i, RSS_a = 150.$$

Модель Б:

$$y_i = \beta_1 + \beta_4 w_i + u_i, RSS_b = 200.$$

Найдите значение F-статистики для проверки гипотезы об истинности ограниченной модели.

171. Рассмотрим две конкурирующие модели: $ARMA(2, 1)$ с ненулевым ожиданием и $ARMA(1, 2)$ с нулевым ожиданием.
Каким способом разумно выбрать наилучшую из них?

a) ADF тест с трендом
b) нет верного ответа
c) критерий Акаике
d) ADF тест с константой
e) KPSS тест с константой
f) ADF тест без константы

172. Процесс случайного блуждания с дрейфом описывается уравнением

   a) $X_t = \mu + u_t$
   b) $X_t = \mu + 0.7X_{t-1} + u_t$
   c) $X_t = X_{t-1} + u_t$
   d) $X_t = 0.7X_{t-1} + u_t$
   e) $X_t = \mu + X_{t-1} + u_t$

173. При наличии сильной практической мультиколлинеарности нарушается следующее свойство МНК-оценок параметров классической регрессии:

   a) эффективность в классе линейных и несмещенных оценок
   b) линейность по зависимой переменой
   c) несмещённость
   d) единственных оценок МНК не существует
   e) равенство нулю суммы остатков

174. Рассмотрим уравнение

$$y_t = 8 + 18y_{t-1} + u_t + 0.3u_{t-1},$$

где (u_t) — белый шум.
Сколько стационарных решений вида $MA(\infty)$ относительно шума (u_t) имеет это уравнение?

   a) 1
   b) 3
   c) бесконечно много
   d) 2
   e) 0
   f) 4

175. Рассмотрим алгоритм случайного леса для задачи регрессии. Каждое дерево даёт свой прогноз. Выберите верное утверждение о механизме агрегирования прогнозов отдельных деревьев в прогноз леса.

   a) вес, с которым прогноз дерева учитывается в прогнозе леса, линейно растёт по номеру дерева
   b) вес, с которым прогноз дерева учитывается в прогнозе леса, нелинейно убывает по номеру дерева
   c) вес, с которым прогноз дерева учитывается в прогнозе леса, нелинейно растёт по номеру дерева
   d) прогноз леса равен среднему арифметическому прогнозов отдельных деревьев
   e) нет верного ответа
   f) вес, с которым прогноз дерева учитывается в прогнозе леса, линейно убывает по номеру дерева

176. В рамках ETS(ANN) модели найдите точечный прогноз на один шаг вперёд, Последнее значение уровня (сглаженного ряда) равно 100, $\ell_t = \ell_{t-1} + 0.3u_t$, $u_t \sim \mathcal{N}(0; 16)$, $y_t = \ell_{t-1} + u_t$.
Ответ введите с точностью до двух знаков после десятичной точки.

177. Винни-Пух построил классический доверительный интервал для коэффициента β_z в модели $y_i = \beta_1 + \beta_x x_i + \beta_z z_i + u_i$. Интервал получился широкий, $[-20; 40]$.
Какую оценку $\hat{\beta}_z$ получил Винни-Пух?

178. Какой период у функции $\cos(2\pi t/365)$?
Ответ укажите с точностью до двух знаков после десятичной точки.

179. Рассмотрим модель ETS(AAdN). Для последнего наблюдения $\ell_{100} = 100$, $b_{100} = 4$, $\phi = 0.5$.
Найдите точечный прогноз для y_{104}. Ответ укажите с точностью до двух знаков после запятой.

180. Агнесса оценила регрессию $\hat{y}_i = 0.7 + 2x_i + 3w_i$ с помощью МНК по 1000 наблюдений. Классические стандартные ошибки коэффициентов равны 0.1, 0.2 и 0.3 соответственно.
Наблюдения являются независимыми, на ошибки выполнены классические предпосылки, $\mathbb{E}(u_i \mid X) = 0$, $\operatorname{Var}(u_i \mid X) = \sigma^2$, $\operatorname{Cov}(u_i, u_j \mid X) = 0$ для $i \neq j$.
Найдите 80% доверительный интервал для коэффициента β_x.
Можно использовать функции из статистических пакетов или таблицы.
В ответе укажите правую границу интервала с точнотью до двух знаков после десятичной точки.

181. Процесс (y_t) стационарный.
Найдите максимальное значение t, при котором гарантированно выполнено условие

$$\operatorname{Cov}(y_{106}, y_{113}) = \operatorname{Cov}(y_{120}, y_t).$$

182. Рассмотрим модель ETS(AAdN). Для последнего наблюдения $\ell_{100} = 100$, $b_{100} = 2$, $\phi = 0.5$.
Найдите точечный прогноз для y_{103}. Ответ укажите с точностью до двух знаков после запятой.

183. Выберите верное утверждение о модели бинарного выбора:

a) значимость коэффициентов проверяется с помощью статистики, имеющей t-распределение

b) нельзя включать в качестве независимых дамми-переменные

c) недостатком линейной вероятностной модели является возможная нереалистичность значений вероятности

d) оценки коэффициентов логит и пробит моделей всегда имеют один и тот же знак

e) ROC кривая является выпуклой для любой логит-модели

184. Рассмотрим временной ряд из 100 наблюдений, оканчивающийся наблюдениями 6 и 5.
Ряд описывается моделью $y_t = y_{t-1} + 2 + u_t$, где величины (u_t) независимы и нормально распределены $\mathcal{N}(0; 4)$.
Постройте 95%-й предиктивный интервал для y_{101}. В ответе укажите правую границу интервала с точностью до двух знаков после запятой.

185. Рассмотрим временной ряд из 100 наблюдений, оканчивающийся наблюдениями 6 и 5.
Ряд описывается моделью $y_t = 6 + u_t$, где величины (u_t) независимы и нормально распределены $\mathcal{N}(0; 4)$.
Постройте 95%-й предиктивный интервал для y_{101}. В ответе укажите правую границу интервала с точностью до двух знаков после запятой.

186. Величины X_1, ..., X_{17} одинаково распределены и независимы. Обозначим их сумму буквой S.
Найдите частную корреляцию $\operatorname{pCorr}(X_1, X_2; S)$.
Ответ вводите с точностью до двух знаков после десятичной точки.

187. По одним и тем же наблюдениям оценили две регрессии: $\hat{y} = 1 + 3X_1$ и $\hat{y} = 2 + 5X_2$.
Известно, что $\widehat{\operatorname{Cov}}(X_1, X_2) > 0$. Оценки МНК коэффициентов регрессии $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$:

a) $\hat{\beta}_0 = 1.5, \hat{\beta}_1 = 3, \hat{\beta}_2 = 5$

b) $\hat{\beta}_1, \hat{\beta}_2$ найти невозможно, $\hat{\beta}_0 = 3$

c) $\hat{\beta}_0 = 3, \hat{\beta}_1 = 3, \hat{\beta}_2 = 5$

d) $\hat{\beta}_0$ найти невозможно, $\hat{\beta}_1 = 3, \hat{\beta}_2 = 5$

e) оценки коэффициентов невозможно найти по имеющимся данным

188. Выберите верную формулу

a) $\sum(y_i - \bar{y})(x_i - \bar{x}) = \sum y_i x_i + \sum \bar{y}\bar{x}$

b) $\sum(y_i - \bar{y})(x_i - \bar{x}) = \sum y_i x_i$

c) $\sum(y_i - \bar{y})(x_i - \bar{x}) = \bar{y}\bar{x}$

d) $\sum(y_i - \bar{y})(x_i - \bar{x}) = \sum y_i(x_i - \bar{x})$

e) $\sum(y_i - \bar{y})(x_i - \bar{x}) = -\sum \bar{y}(x_i - \bar{x})$

189. Оценки коэффициентов линейной регрессии, полученные методом максимального правдоподобия и методом наименьших квадратов в случае нормально распределенной случайной составляющей, будут совпадать

a) если ковариационная матрица случайной составляющей нулевая

b) если ковариационная матрица случайной составляющей пропорциональна единичной

c) если ковариационная матрица случайной составляющей диагональна

d) никогда

e) всегда

190. Показатель R^2_{adj} можно вычислить по формуле

a) $R^2_{adj} = \frac{k-1}{n-k} - R^2 \cdot \frac{n-1}{n-k}$

b) $R^2_{adj} = \frac{k-1}{n-k} + R^2 \cdot \frac{n-k}{n-1}$

c) $R^2_{adj} = \frac{k-1}{n-k} + R^2 \cdot \frac{n-1}{n-k}$

d) $R^2_{adj} = \frac{n-k}{k-1} + R^2 \cdot \frac{n-1}{n-k}$

e) $R^2_{adj} = (-1) \cdot \frac{k-1}{n-k} + R^2 \cdot \frac{n-1}{n-k}$

f) $R^2_{adj} = \frac{k}{n-k} + R^2 \cdot \frac{n-1}{n-k}$

191. Найдите наибольший корень характеристического уравнения AR-части рекуррентного уравнения

$$y_t = y_{t-1} + 4y_{t-2} + u_t.$$

Ответ вводите с точностью до двух знаков после десятичной точки.

192. 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 регрессии будут несмещенными и эффективными

c) коэффициент детерминации R^2 окажется отрицательным

d) МНК-оценки параметров окажутся неэффективными в классе линейных и несмещённых

e) МНК-оценки параметров окажутся смещёнными

193. Если $\alpha = 0.05$ и P-значение равно 0.04, то

a) H_a отвергается

b) H_0 принимается

c) недостаточно информации для ответа

d) H_a принимается

e) H_a не отвергается

f) H_0 отвергается

194. Выберите верное утверждение про штрафную функцию LASSO.

a) Растёт при росте RSS, но падает по мере удаление оценок коэффициентов от нуля.

b) Не зависит от RSS, падает по мере удаление оценок коэффициентов от нуля.

c) Растёт при росте RSS и по мере удаление оценок коэффициентов от нуля.

d) Падает при росте RSS, но растёт по мере удаление оценок коэффициентов от нуля.

e) Растёт при росте RSS, не зависит от удалённости оценок коэффициентов от нуля.

f) Падает при росте RSS и по мере удаление оценок коэффициентов от нуля.

195. Рассмотрим уравнение

$$y_t = 7 + 15y_{t-1} + u_t + 0.3u_{t-1},$$

где (u_t) — белый шум.
Сколько стационарных решений имеет это уравнение?

a) 1

b) 3

c) 4

d) бесконечно много

e) 2

f) 0

196. Известно, что $y_1 = 7$, $y_2 = 10$, а остальные наблюдения — пропущенные. Для восстановления наблюдений Винни-Пух использует модель $\mathbb{E}(y_t \mid y_{t-1}, \ldots, y_1) = 4 + 0.4y_{t-1}$.
Помогите Винни-Пуху восстановить y_4 с точностью до двух знаков после десятичной точки.

197. Исследователь Пантелеймон оценивает парную регрессию $y_i = \beta_1 + \beta_2 x_i + u_i$. На ошибки выполнены классические предпосылки, $\mathbb{E}(u_i \mid X) = 0$, $\mathrm{Var}(u_i \mid X) = \sigma^2$, $\mathrm{Cov}(u_i, u_j \mid X) = 0$ для $i \neq j$.
Пантелеймон знает, что сумма квадратов остатков в исходной регрессии оказалась равной 100, а во вспомогательной регрессии $\hat{x}_i = \hat{\gamma}$ сумма квадратов остатков равна 20. Обе регрессии построены по 102 наблюдениям.
Какую несмещённую оценку для $\mathrm{Var}(\hat{\beta}_2 \mid X)$ получит Пантелеймон?
Ответ вводите с точностью до двух знаков после десятичной точки.

198. Свободно распространяемым программным обеспечением является

a) SPSS

b) R

c) Stata

d) Matlab

e) Excel

f) Eviews

199. В модели 10 параметров и функция правдоподобия в точке максимума равна 10^{-10}.
Найдите значение критерия Акаике с точностью до двух знаков после десятичной точки.

200. К несостоятельности МНК-оценок вектора коэффициентов приводит

a) корреляция ошибок по схеме AR(1)

b) условная гетероскедастичность ошибок

c) эндогенность одного из регрессоров

d) корреляция между регрессорами

e) корреляция ошибок по схеме MA(1)

f) нестрогая мультиколлинеарность

201. Рассмотрим $MA(1)$ модель $y_t = u_t + 8u_{t-1}$, где (u_t) — белый шум с дисперсией 7.
Найдите долгосрочную дисперсию процесса (y_t).

202. Рассмотрим $MA(1)$ модель $y_t = u_t + 7u_{t-1}$, где (u_t) — белый шум с дисперсией 7.
Найдите долгосрочную дисперсию процесса (y_t).

203. Если в регрессии обнаружена автокорреляция типа AR(1), то статистика Дарбина-Уотсона и оценка коэффициента автокорреляции $\hat{\rho}$ связаны между собой соотношением

a) $\hat{\rho} \approx 2(1 - DW)$
b) $DW \approx 2(1 - \hat{\rho})$
c) $DW \approx \hat{\rho}$
d) $DW \approx \hat{\rho}/2$
e) $\hat{\rho} \approx DW/2$

204. Запись $3.20E - 16$ означает

a) $3.2 \cdot e - 16$
b) Ошибка с кодом 16
c) 3.2^{e-16}
d) $3.2 \cdot e^{-16}$
e) $3.2 \cdot (e - 16)$
f) $3.2 \cdot 10^{-16}$

205. Если в уравнение регрессии не включена константа, то

a) R^2 является показателем качества подгонки регрессии
b) К этой модели применима теорема Гаусса-Маркова
c) Сумма остатков регрессии равна 0
d) R^2_{adj} в этой модели всегда неотрицательный
e) Значимость коэффициентов регрессии нельзя проверять при помощи t-статистики

206. Если все Y_i в линейной регрессии увеличить в два раза, то оценка $\hat{\beta}_2$

a) изменится в произвольную сторону, в зависимости от X_i
b) поделится на 4
c) не изменится
d) помножится на 4
e) поделится на 2
f) помножится на 2

207. В модели парной линейной регрессии со свободным членом $Y_i = \beta_0 + \beta_1 X_i + u_i$ несмещённой оценкой дисперсии оценки МНК $\hat{\beta}_1$ является

a) $RSS/(n-2)$
b) $\sum(Y_i - \bar{Y})^2/(n-2)$
c) $\sum(Y_i - \bar{Y})^2/(n-1)$
d) $RSS/((n-2)\sum_i (X_i - \bar{X})^2)$
e) RSS/n

208. В модели множественной регрессии $\hat{y}_i = 2 + 3x_i + 7z_i - 6w_i$ оценённая сумма квадратов равна 777. Для оценивания использовали 200 наблюдений.
Найдите сумму $\sum_i \hat{y}_i(\hat{y}_i - \bar{y})$.

209. Рассмотрим уравнение

$$y_t = 8y_{t-1} + 4y_{t-2} + u_t,$$

где (u_t) — белый шум.
Сколько нестационарных решений оно имеет?

a) нет верного ответа
b) 4
c) бесконечность
d) 3
e) 2
f) 1

210. Исследователь Иа-Иа использует аддитивное разложение ряда. Очередное наблюдение равно 60, трендовая составляющая равна 40, сезонная составляющая 30.
Найдите остаточную компоненту.

211. Пантелеймон оценил парную регрессию по 50 наблюдениям

$$\hat{y}_i = 8.9 + 12.9x_i.$$

Оценка дисперсии прогноза при $x = 1$ равна 100. Оценка дисперсии случайной ошибки равна 400. Предпосылки классической линейной модели выполнены, ошибки имеют нормальное распределение.
Постройте 95%-й доверительный интервал для $\mathbb{E}(y_i \mid x_i = 1)$.
В ответе укажите правую границу интервала с точностью до двух знаков после десятичной точки.

212. По набору данных cars из R оцените модель

$$dist_i = \beta_1 + \beta_2 speed_i + u_i.$$

Ошибки модели нормально распределены и удовлетворяют классическим предпосылкам.
Для машины со стартовой скоростью 20 миль в час постройте 95%-й доверительный интервал для ожидаемой длины тормозного пути в футах.
В ответе укажите правую границу интервала с точностью до двух знаков после десятичной точки.
Набор данных встроен в R и доступен по ссылке: https://github.com/vincentarelbundock/Rdatasets/raw/maste

213. Методом максимального правдоподобия Гоша оценил модель

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

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

a) 2.4
b) 0.48
c) 0.5
d) $24/7$
e) не существует
f) 2

214. Оценка ковариационной матрицы оценок коэффициентов $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ имеет вид

$$\begin{pmatrix} 5 & 1 & 1 \\ 1 & 6 & -1 \\ 1 & -1 & 9 \end{pmatrix}.$$

Найдите стандартную ошибку $se(\hat{\beta}_2 - \hat{\beta}_3)$.
Ответ вводите с точностью до двух знаков после десятичной точки.

215. Найдите наименьшее возможное значение q, если известно, что $SARIMA(3,0,3)(1,0,1)[12]$-модель является частным случаем $ARMA(p,q)$-модели.

216. Ряд (u_t) — белый шум с дисперсией равной 16.
Найдите ковариацию $\mathrm{Cov}(u_t + 2u_{t-1} + 4u_{t-2}, u_{t-1} + 2u_{t-2} + 7u_{t-3})$.

217. Инструмент Z_t для оценивания динамической модели $Y_t = \beta_1 + \beta_2 X_t + \beta_3 Y_{t-1} + u_t$ с экзогенным вектором X и AR(1) процессом в ошибках u_t должен удовлетворять требованию

a) $\mathrm{Corr}(Y_{t-1}, Z_t) = 0$
b) $\mathrm{Corr}(u_t, Z_t) \to 1$
c) $\mathrm{Corr}(X_t, Z_t) = 0$
d) $\mathrm{Corr}(u_t, Z_t) = 0$
e) $\mathrm{Corr}(Y_{t-1}, Z_t) \neq 0$

218. В модели $Y_i = \beta_0 + \beta_1 X_i + u_i$ при выполненных предпосылках теоремы Гаусса-Маркова и нормальных ошибках тестовая статистика $(\hat{\beta}_1 - \beta_1)/se(\hat{\beta}_1)$ имеет распределение

a) $\mathcal{N}(0;\sigma^2)$
b) χ^2_1
c) χ^2_{n-2}
d) $\mathcal{N}(0;1)$
e) t_{n-2}

219. Рассмотрим метод главных компонент для переменных с разными единицами измерения. Выберите верное утверждение.

a) Обязательно нужно привести переменные к общему масштабу.
b) Важно убедиться в некоррелированности исходных переменных.
c) Необходимо и достаточно центрировать исходные переменные.
d) Важно убедиться в ортогональности исходных переменных.
e) Обязательно нужно сохранить исходные единицы измерения переменных.

220. Рассеянная исследовательница Надежда ошибочно оценила модель $y_i = \beta_1 + \beta_x x_i + \beta_w w_i + u_i$ и получила неверные оценки коэффициентов $\hat{\beta}_{wrong}$. Корректно используя неверные оценки коэффициентов, она нашла остатки $\hat{u}_{wrong}$ и прогнозы $\hat{y}_{wrong}$.
Верные величины обозначим $\hat{\beta}$, $\hat{u}$ и $\hat{y}$.
Выберите верное утверждение.

a) $\hat{u}_{wrong} \perp \hat{y}_{wrong}$
b) $\hat{\beta} \perp \hat{\beta}_{wrong}$
c) $\hat{u} \perp \hat{\beta}_{wrong}$
d) $\hat{u}_{wrong} \perp \hat{y}$
e) $\hat{u} \perp \hat{y}_{wrong}$
f) $\hat{u}_{wrong} \perp \hat{\beta}$

221. С какой целью при оценивании модели $MA(2)$ накладывают требование обратимости процесса?

a) чтобы гарантировать стационарность процесса
b) чтобы обратить характеристический многочлен
c) чтобы обратить лаговый многочлен
d) чтобы гарантировать неотрицательность дисперсии процесса
e) чтобы гарантировать единственность оценок
f) чтобы гарантировать нулевые корреляции высоких порядков

222. Оценена зависимость расходов потребителей на газ и электричество Y в США в 1977-1999 г. в постоянных ценах I квартала 1977 г. от времени ($t = 1$ для 1977, $t = 2$ для 1978 и т.д.) с учётом сезонных факторов ($D_i = 1$, если наблюдение относится к i-ому кварталу и 0 иначе, $i = 1, \ldots, 4$):
$\hat{y} = 8 + 0.1t - 3D_2 - 2.6D_3 - 2D_4$
Если в качестве базовой категории будет принят не первый квартал, а третий, уравнение регрессии примет вид

a) $\hat{y} = 8 + 0.1t + 3D_1 + 2.6D_2 + 2D_4$
b) $\hat{y} = 5.4 + 0.1t - 3D_1 - 2.6D_2 - 2D_4$
c) $\hat{y} = 5.4 + 0.1t - 3D_1 - 0.4D_2 - D_4$
d) $\hat{y} = 5.4 + 0.1t + 2.6D_1 - 0.4D_2 + 0.6D_4$
e) $\hat{y} = 8 + 0.1t - 3D_1 - 2.6D_2 - 2D_4$

223. Рассмотрим модель

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i.$$

Все предпосылки классической регрессионной модели выполнены. По 25 наблюдениям оказалось, что сумма квадратов остатков равна $RSS = 60$.
Найдите несмещённую оценку для дисперсии случайной ошибки.
Ответ вводите с точность до двух знаков после десятичной точки.

224. По 52 наблюдениям студент построил две регрессии, $\hat{Y}_i = 3.1 + 0.8X_i$ и $\hat{X}_i = -0.3 + 0.2Y_i$. Коэффициент R^2_{adj} для первой регрессии примерно равен

a) 0.40
b) 0.14
c) 0.16
d) 0.32
e) 0.37

225. Для выбора между линейной и полулогарифмической моделями (где EARNINGS — почасовая заработная плата в $, S — длительность обучения, ASVABC - результаты тестов, характеризующие успеваемость) был проведен тест Дэвидсона, Уайта и МакКиннона и получены следующие результаты:

	Зависимая: Y	Зависимая: $\ln Y$
(Intercept)	-26.148	-1.941
	(4.17)	(3.2499)
S	2.008	0.087
	(0.276)	(0.035)
ASVABC	0.393	0.017
	(0.079)	(0.007)
lin_add	-15.373	
	(5.984)	
semilog_add		-0.029
		(0.065)
R^2	0.2071	0.2212
F	46.59	50.74
Adj. R^2	0.2027	0.2168
Num. obs.	540	540
RSS	90975.57	148.1
$\hat{\sigma}$	13.04	0.5256

Где lin_add $= \ln(\hat{y}) - \widehat{\ln Y}$, semilog_add $= \hat{y} - \exp(\widehat{\ln Y})$ и в скобках указаны стандартные ошибки.
На уровне значимости 5% можно сделать вывод, что

a) Лучше линейная модель
b) Между линейной и полулогарифмической моделями нет статистической разницы
c) Лучше линейная в логарифмах модель
d) Невозможно выбрать лучшую модель
e) Лучше полулогарифмическая модель

226. Предпосылки теоремы Гаусса-Маркова выполнены, случайные ошибки нормально распределены, уровень доверия равен 80%, критическое значение t-статистики равно 1.64, всего n наблюдений. Регрессия имеет вид $\hat{Y}_i = \underset{(3)}{-4} + \underset{(0.2)}{5} X_i$, в скобках указаны стандартные ошибки. Доверительный интервал для β_2 равен

    a) $[1.72; 8.28]$
    b) не существует
    c) $[4.27; 5.73]$
    d) $[4.67; 5.33]$
    e) $[4.85; 5.15]$
    f) $[3.36; 6.64]$

227. Илон Маск по 100 наблюдениям оценил множественную регрессию с константой с помощью МНК. Оказалось, что $\sum(\hat{y}_i - \bar{y})^2 = 200$, и $\sum(y_i - \bar{y})^2 = 600$.
Чему равна средняя величина квадрата остатка?

228. Гипотеза Алмон при оценивании модели с распределенными лагами $Y_t = \alpha + \sum_{j=0}^{J} \beta_j X_{t-j} + u_t$ состоит в том, что коэффициент β_j представим в виде

    a) $\alpha_0+\alpha_1\lambda+\alpha_2\lambda^2+\ldots+\alpha_j\lambda^j$
    b) $\alpha_0 \ln j$
    c) $\alpha_0+\alpha_1 j+\alpha_2 j^2+\ldots+\alpha_r j^r$
    d) $\alpha_0\lambda^j$
    e) $\alpha_0 + \alpha_1 j$

229. Рассмотрим модель $Y = X\beta + u$. Условия теоремы Гаусса-Маркова выполнены, причём $\mathrm{Var}(u_i) = \sigma_u^2$, $\hat{y} = PY$, $P = X(X'X)^{-1}X'$ и I - единичная матрица. Ковариационная матрица случайного вектора $e = Y - \hat{y}$ равна

    a) $\sigma_u^2(I+P)$
    b) $\sigma_u^2(P-I)$
    c) $\sigma_u^2 P$
    d) $\sigma_u^2(I-P)$
    e) $\sigma_u^2 I$

230. В исходной выборке 10 наблюдений. Найдите ожидаемое количество копий второго наблюдения в очередной бутстрэп выборке.
Ответ укажите с точностью до двух знаков после десятичной точки.

231. Логарифм наблюдаемой величины описывается ETS(AAN) моделью, $\ln y_t \sim ETS(AAN)$. Обычный 95%-й предиктивный интервал для $\ln y_{T+1}$ имеет вид $[10; 20]$.
Найдите условное ожидание y_{T+1} с точностью до целых.

232. Петя и Вася проверяют гипотезу $\beta_x = 0$ против альтернативной $\beta_x \neq 0$ по одним и тем же данным, одни и тем же способом. Единственная разница в том, что Петя использует уровень значимости $\alpha = 0.01$, а Вася — 0.02.
Рассмотрим 4 ситуации:
А. Петя отверг H_0, Вася отверг H_0.
Б. Петя отверг H_0, Вася не отверг H_0.
В. Петя не отверг H_0, Вася отверг H_0.
Г. Петя не отверг H_0, Вася не отверг H_0.
Какие из этих ситуаций возможны?

    a) только В, Г
    b) только А и Г
    c) все возможны
    d) только А, Б, Г
    e) только А, Б
    f) только А, В и Г

233. Исследователь выполнил второй шаг в PE-тесте МакКиннона. В регрессии $\ln Y_i$ на исходные регрессоры и $Z_i = \hat{Y}_i - \exp(\widehat{\ln Y_i})$ коэффициент при Z_i оказался значимым. А в регрессии Y_i на исходные регрессоры и $W_i = \ln \hat{Y}_i - \widehat{\ln Y_i}$ коэффициент при W_i оказался незначимым. Из результатов следует сделать вывод, что

a) следует предпочесть линейную модель
b) в исходной модели пропущен регрессор Z_i
c) следует предпочесть логарифмическую модель
d) в исходной модели пропущен регрессор W_i
e) тесты противоречат друг другу, ни одна из моделей не предпочитается
f) следует предпочесть полулогарифмеческую модель

234. Частным случаем какой модели является модель за тета-методом?

a) ETS(MMM) с $\alpha = 1$
b) нет верного ответа
c) ETS(AAN) с $\beta = 0$
d) ETS(AAA) с $\alpha = 0$
e) ETS(MMM) с $\beta = 1$
f) ETS(AAA) с $\gamma = 0$

235. Найдите наибольший корень характеристического уравнения AR-части рекуррентного уравнения

$$y_t = y_{t-1} + 10y_{t-2} + u_t.$$

Ответ вводите с точностью до двух знаков после десятичной точки.

236. Если основная гипотеза в тесте Дики-Фуллера отвергается, то временной ряд является

a) стационарным в первых разностях
b) нестационарным
c) нормально распределённым
d) коинтегрированным
e) стационарным

237. Выборочная дисперсия остаточной компоненты ряда равна 100, выборочная дисперсия суммы тренда и остаточной компоненты равна 90.
Найдите силу выраженности тренда с точностью до двух знаков после десятичной точки.

238. Спящая Красавица построила парную регрессию с $R^2 = 0.9$ по 99 наблюдениям. Злая Фея-Крестная Карабос добавляет Спящей Красавице ровно одно наблюдение, чтобы максимально снизить коэффициент детерминации.
Какого значения R^2 сможет добиться Злая Фея?

239. Выберите верное равенство для множественной регрессии $y_i = \beta_x x_i + \beta z z_i + \beta_w w_i + u_i$, оцениваемой с помощью МНК.
Остатки обозначим $\hat{u}_i$.

a) $\sum \hat{u}_i = 0$
b) $\sum x_i u_i = 0$
c) $\sum \hat{u}_i y_i = 0$
d) $\sum x_i y_i = 0$
e) $\sum x_i z_i = 0$
f) $\sum w_i \hat{u}_i = 0$

240. Рассмотрим алгоритм случайного леса для задачи регрессии.
Какую целевую переменную «учится» прогнозировать дерево номер 4?

a) разницу между исходной переменной y_t и суммой прогнозов первых трёх деревьев

b) нет верного ответа

c) разницу между исходной переменной y_t и суммой прогнозов первых трёх деревьев, домноженной на темп обучения

d) разницу между исходной переменной y_t и прогнозом третьего дерева

e) исходную переменную y_t

f) разницу между исходной переменной y_t и прогнозом третьего дерева, домноженным на темп обучения

241. Величины X_1, ..., X_{20} одинаково распределены и независимы. Обозначим их сумму буквой S. Найдите частную корреляцию $\mathrm{pCorr}(X_1, X_2; S)$.
Ответ вводите с точностью до двух знаков после десятичной точки.

242. Для оцениваемой по 30 наблюдениям регрессии $Y_i = \alpha + \beta X_i + u_i, i = 1, \ldots, n$ известны суммы $\sum_{i=1}^{30} X_i = -15, \sum_{i=1}^{30} X_i^2 = 60, \sum_{i=1}^{30} X_i Y_i = 15, \sum_{i=1}^{30} Y_i = 75$. Система нормальных уравнений для оценок коэффициентов регрессии α, β методом наименьших квадратов равносильна системе

a) $4\alpha - 6\beta = 1; 6\alpha + 60\beta = 75$

b) $2\alpha - \beta = 5; \alpha - 4\beta = -1$

c) $30\alpha + 15\beta = 75; 15\alpha + 60\beta = 15$

d) $2\alpha - \beta = -1; \alpha - 4\beta = 5$

e) $30\alpha - 15\beta = 15; -15\alpha - 12\beta = 1$

243. Два временных ряда могут быть коинтегрированными, только если

a) один ряд стационарен, а второй — нет

b) оба ряда нестационарны и имеют разные порядки интегрирования

c) оба ряда стационарны вокруг тренда

d) оба ряда нестационарны и имеют одинаковый порядок интегрирования

e) оба ряда стационарны

244. Случайные величины r и s независимы и равновероятно равны 0 или 1. Чему равно условное ожидание $\mathbb{E}(rs + s^2 \mid r)$?

a) $r + 0.5$

b) $0.5r + 0.5$

c) $rs + s^2$

d) $0.5r + 0.25$

e) $r + 1$

245. Укажите число параметров в $ARIMA(4, 1, 1)$ модели с ненулевым математическим ожиданием и нормально распределенными ошибками.

246. Совместное распределение случайных величин X и Y задано с помощью таблицы

	$X = 3$	$X = 4$	$X = 5$
$Y = 3$	0.1	0.3	0.1
$Y = 4$	0.15	0.05	0.05
$Y = 6$	0.05	0.15	0.05

Математическое ожидание случайной величины Y при условии, что $X = 3$, равно

a) 2

b) 3.4

c) 6

d) 4

e) 2.4

247. Рассмотрим метод максимального правдоподобия для оценки одного параметра. Как можно получить оценку дисперсии оценки параметра?

a) Посчитать вторую производную лог-правдоподобия со знаком минус.

b) Обратить вторую производную лог-правдоподобия.

c) Обратить первую производную лог-правдоподобия.

d) Обратить вторую производную лог-правдоподобия со знаком минус.

e) Обратить первую производную лог-правдоподобия со знаком минус.

f) Посчитать первую производную лог-правдоподобия со знаком минус.

248. Если в регрессии отсутствует свободный член, то в общем случае

a) $\sum_i \hat{Y}_i = \sum_i Y_i$

b) сумма остатков равна нулю

c) $TSS \neq ESS + RSS$

d) сумма квадратов остатков равна нулю

e) R^2 является мерой качества подгонки регрессии

249. Портос построил регрессию по 66 наблюдениям, $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 W_i + \hat{\beta}_3 Z_i$, $RSS = 140$. Затем Портос оценил вспомогательную регрессию, $\hat{\hat{Y}}_i = \hat{\gamma}_0 + \hat{\gamma}_1 X_i + \hat{\gamma}_2 W_i + \hat{\gamma}_3 Z_i + \hat{\delta}_2 \hat{Y}_i^2 + \hat{\delta}_3 \hat{Y}_i^3$, $RSS = 120$. При проверке гипотезы о правильной спецификации модели в тесте Рамсея F-статистика равна

a) 6

b) $30/7$

c) $11/3$

d) $10/3$

e) 5

250. Условие порядка для любого уравнения из системы может быть сформулировано следующим образом. Число эндогенных переменных, включенных в уравнение, уменьшенное на 1, должно быть

a) не меньше числа экзогенных переменных, включенных в это уравнение

b) не меньше числа экзогенных переменных, исключенных из этого уравнения

c) не больше числа экзогенных переменных, включенных в это уравнение

d) не больше числа эндогенных переменных, включенных в это уравнение

e) не больше числа эндогенных переменных, исключенных из этого уравнения

f) не больше числа экзогенных переменных, исключенных из этого уравнения

251. Выберите верное утверждение про ARDL модель с зависимой переменной y_t, предиктором x_t и ошибкой u_t.

a) u_t является $ARMA(p,q)$ процессом с $p \geq 1$ и $q \geq 1$

b) u_t является $AR(p)$ процессом с $p \geq 1$

c) u_t является $MA(q)$ процессом с $q \geq 1$

d) Δy_t является нестационарным

e) нет верного ответа

f) u_t является белый шумом

252. Сколько свободных параметров оценивается в $ETS(AAA)$ модели по квартальным данным?

253. По исходному временному ряду была оценена множественная регрессия

$$\hat{y}_t = 3 + 0.6y_{t-1} - 0.09y_{t-2}.$$

Найдите второе значение выборочной частной автокорреляционной функции $PACF_2$ с точностью до двух знаков после десятичной точки.

254. Исследователь Ярополк хочет оценить коэффициент β_x в модели $y_i = \beta_1 + \beta_x x_i + u_i$.
По исходным наблюдениям он рассчитал суммы $\sum(x_i - \bar{x})(y_i - \bar{y}) = 300$, $\sum(y_i - \bar{y})^2 = 1000$ и $\sum(x_i - \bar{x})^2 = 100$.
Чему равна МНК-оценка $\hat{\beta}_x$?

255. По данным для 27 фирм исследована зависимость прибыли Y от числа работников X вида\ $Y = \beta_0 + \beta_1 X + u$ и получено $\hat{\beta}_0 = 8, \hat{\beta}_1 = 2, \hat{\sigma}^2 = 25$ и матрица $(X'X)^{-1} = \begin{pmatrix} 0.36 & -0.03 \\ -0.03 & 0.09 \end{pmatrix}$.\ 95% доверительный интервал для β_1:

a) [-1.09; 5.09]
b) невозможно вычислить по имеющимся данным
c) [0.04; 3.96]
d) [-0.94; 4.94]
e) [1.82; 14.18]

256. Что может произойти с R^2 при добавлении нового наблюдения в модель множественной регрессии с константой?

a) Коэффициент R^2 может остаться неизменным или упасть.
b) Коэффициент R^2 может остаться неизменным или вырасти.
c) Коэффициент R^2 обязательно упадёт.
d) Коэффициент R^2 обязательно вырастет.
e) Коэффициент R^2 может измениться в любую сторону.

257. Корни лагового многочлена AR-части уравнения равны $\ell_1 = 12$ и $\ell_2 = 28$.
Найдите наибольший корень характеристического уравнения AR-части.
Ответ вводите с точностью до двух знаков после десятичной точки.

258. В предположениях нормальности ошибок ширина 95%-го интервала для ожидаемого (среднего) значения Y_{n+1} равна 1200. Известно, что $\hat{\sigma} = 400$ и $n = 60$. Ширина 95%-го интервала для фактического (индивидуального) значения Y_{n+1} примерно равна

a) 1000
b) 2000
c) 1500
d) 1400
e) 1600

259. Рассмотрим модель $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i$. Гипотезу

$$\begin{cases} \beta_2 + \beta_3 = 1 \\ \beta_5 = 0 \end{cases}$$

можно проверить с помощью оценки дополнительной модели

a) $Y_i - X_{i3} = \beta_1 + \beta_2(X_{i2} + X_{i3}) + \beta_4 X_{i4} + u_i$
b) $Y_i - \beta_2 = \beta_1 + \beta_2(X_{i2} + X_{i3}) + \beta_4 X_{i4} + u_i$
c) $Y_i - X_{i3} = \beta_1 + \beta_2(X_{i2} - X_{i3}) + \beta_4 X_{i4} + u_i$
d) $Y_i = \beta_1 + \beta_2(X_{i2} + X_{i3} - 1) + \beta_4 X_{i4} + u_i$
e) $Y_i - X_{i2} = \beta_1 + \beta_2(X_{i2} - X_{i3}) + \beta_4 X_{i4} + u_i$

260. Рассмотрим короткий ряд из четырех наблюдений: 5, 4, 6, 5.
Предположим, что ряд описывается моделью $y_t = y_{t-1} + \mu + u_t$, где величины (u_t) независимы и нормально распределены $\mathcal{N}(0;\sigma^2)$.
Трактуя y_1 как фиксированную константу, найдите оценку $\hat{\sigma}^2$ методом максимального правдоподобия с точностью до двух знаков после десятичной точки.

261. Распределение случайной величины X задано таблицей

x	0	1	2	3
$\mathbb{P}(X=x)$	$-b$	$0.5-b$	$0.5+b$	b

Вероятность $\mathbb{P}(X=1)$ равна

a) 0.5
b) 0.3
c) 0.2
d) 0
e) 0.4

262. Какая зависимость математического ожидания исходного процесса от времени предполагается в нулевой гипотезе $KPSS$-теста с константой?

a) $\mathbb{E}(y_t) = \mu + \alpha t$
b) $\mathbb{E}(y_t)$ строго монотонна
c) нет верного ответа
d) $\mathbb{E}(y_t) = \mu + \alpha t + \beta t^2$
e) $\mathbb{E}(y_t)$ строго возрастает
f) $\mathbb{E}(y_t) = \mu$

263. Исследователь Леонардо оценил модель логистической регрессии $\hat{\mathbb{P}}(y_i = 1) = \Lambda(0.4 - 0.2x_i)$. Найдите по этой модели прогноз вероятности для $x_i = 2$.

a) $\exp(0.5)$
b) 0.75
c) 1
d) 0.25
e) 0
f) 0.5

264. Переменная Y_i принимает значения 0 или 1. Логарифмическая функция правдоподобия, используемая для оценивания логит и пробит моделей, имеет вид

a) $\ln L = \sum_{i=1}^n Y_i \ln F(X_i\beta) - (1-Y_i)\ln(1-F(X_i\beta))$
b) $\ln L = \sum_{i=1}^n Y_i \ln F(X_i\beta) + (1-Y_i)\ln(1-F(X_i\beta))$
c) $\ln L = \sum_{i=1}^n Y_i \ln(1-F(X_i\beta)) + (1-Y_i)\ln F(X_i\beta)$
d) $\ln L = \sum_{i=1}^n Y_i \ln F(X_i\beta) - (1-Y_i)\ln F(X_i\beta)$
e) $\ln L = \sum_{i=1}^n Y_i \ln F(X_i\beta) \cdot (1-Y_i)\ln(1-F(X_i\beta))$

265. Эмманюэль и Владимир оценили зависимость стоимости подержанных Пежо (одной серии) Y от пробега X (измеряемого в км) с помощью модели парной регрессии $Y = \alpha + \beta X + u$ по по одной и той же выборке, однако Эмманюэль измерял стоимость машин в евро, а Владимир – в рублях, 1 евро = 65 рублей. Оценки МНК коэффициента наклона регрессии, полученные Эмманюэлем β_E и Владимиром β_B связаны следующим образом:

a) $\hat{\beta}_B = \hat{\beta}_E$
b) $\hat{\beta}_B = 4225\hat{\beta}_E$
c) $\hat{\beta}_E = 65\hat{\beta}_B$
d) $\hat{\beta}_B = 65\hat{\beta}_E$
e) $\hat{\beta}_E = 4225\hat{\beta}_B$

266. По данным 500 индивидуумов оценили зависимость веса индивидуума Y, измеряемого в фунтах (1 фунт $\approx$ 0.5 кг) от его роста , измеряемого в в футах (1 фут $\approx$ 30 см): $\widehat{\ln Y}_i = -4.4 + 2.4 \ln X_i$. Если рост индивидуума будет измерен в метрах, а вес — в килограммах, то коэффициент перед $\ln X$ будет равен

a) 4.8
b) 1.2
c) 2.4
d) 8

267. Какое условие НЕ требуется в теореме Гаусса-Маркова?

a) случайные ошибки u_i нормально распределены
b) случайные ошибки u_i не коррелированы
c) случайные ошибки u_i имеют одинаковые дисперсии
d) матрица регрессоров X имеет полный ранг
e) модель $Y = X\beta + u$ правильно специфицирована

268. Винни-Пух оценил модель множественной регрессии с константой с помощью МНК. Выборочная корреляция между прогнозами $\hat{y}$ и фактическим вектором y оказалась равной 0.7.
Чему равно отношение $\sum(\hat{y}_i - \bar{y})^2 / \sum(y_i - \bar{y})^2$?

269. У Винни-Пуха есть четыре наблюдения 5, 6, 8, 9 и он оценивает модель вида $y_i = 1 + \beta + u_i$ с помощью МНК.
Оценка $\hat{\beta}$ окажется равна

a) 7
b) 8
c) 9
d) 5
e) 6
f) 10

270. Обобщенный МНК служит для оценивания регрессионной модели $Y = X\beta + u$ в случае нарушения следующего условия теоремы Гаусса-Маркова

a) $\mathrm{Cov}(u_i, X_i) = 0$
b) $\mathrm{Cov}(Y_i, u_i) = 0$
c) $\mathbb{E}(u_i) = 0$
d) $\mathrm{Var}(u_i) = \sigma_u^2$
e) $\mathrm{rank} X = k$

271. По исходному временному ряду была оценена парная регрессия

$$\hat{y}_t = 3 + 0.06 y_{t-2}.$$

Найдите второе значение выборочной автокорреляционной функции ACF_2 с точностью до двух знаков после десятичной точки.

272. Какая зависимость математического ожидания исходного процесса от времени предполагается в альтернативной гипотезе $KPSS$-теста с константой?

a) $\mathbb{E}(y_t) = \mu + \alpha t$
b) нет верного ответа
c) $\mathbb{E}(y_t)$ строго возрастает
d) $\mathbb{E}(y_t)$ строго монотонна
e) $\mathbb{E}(y_t) = \mu + \alpha t + \beta t^2$
f) $\mathbb{E}(y_t) = \mu$

273. На третьем шаге теста МакАлера исследователь получил следующие результаты:

$$\widehat{\ln Y_i} = \underset{(0.20)}{1.54} - \underset{(0.05)}{1.2}\, X_i + \underset{(0.3)}{3.1}\, \hat{v}_{\exp(\widehat{\ln Y}),i}$$

$$\hat{Y}_i = \underset{(0.02)}{1.24} + \underset{(0.03)}{1.1}\, X_i + \underset{(0.2)}{2.4}\, \hat{v}_{\ln \hat{Y},i}$$

На основании этих результатов исследователь

a) должен сделать вывод об отсутствии пропущенных переменных

b) должен предпочесть линейную модель

c) не может выбрать ни логарифмическую, ни линейную модель

d) должен отвергнуть гипотезу об адекватности исходной модели

e) должен предпочесть полулогарифмическую модель

f) должен сделать вывод о наличии пропущенных переменных

274. Выберите верное утверждение про ARDL модель с зависимой переменной y_t и предиктором x_t.

a) y_t является $AR(p)$ процессом с $p \geq 1$

b) y_t является $ARMA(p, q)$ процессом с $p \geq 1$ и $q \geq 1$

c) y_t является $MA(q)$ процессом с $q \geq 1$

d) y_t является нестационарным

e) y_t является стационарным

f) нет верного ответа

275. Если для регрессора используется преобразование Бокса-Кокса с параметром $\theta = -1$, а для зависимой переменной — с параметром $\lambda = 1$, то регрессионное уравнение представимо в виде

a) $\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i$

b) $\ln Y_i = \beta_1 + \beta_2 X_i + u_i$

c) $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$

d) $Y_i = \beta_1 + \beta_2 X_i + u_i$

e) $Y_i = \beta_1 + \beta_2 \frac{1}{X_i} + u_i$

f) $\ln Y_i = \beta_1 - \beta_2 \ln X_i + u_i$

276. С помощью МНК оценена зависимость потребления Y_i от дохода X_i, $\hat{Y}_i = 0.5 - 0.3X_i$. Если же использовать центрированные и нормированные переменные, то зависимость примет вид $\hat{Y}_i^{st} = -0.7X_i^{st}$. Коэффициент множественной детерминации R^2 для первой модели равен

a) 0.3

b) 0.49

c) 0.21

d) 0.7

e) 0.09

277. В рамках ETS(AAN) модели постройте 95%-й предиктивный интервал на два шага вперёд, Последнее значение сглаженного ряда (тренда) равно 100, последний наклон сглаженного ряда (тренда) равен 5, $\alpha = \beta = 0.3$, $u_t \sim \mathcal{N}(0; 16)$.
В ответ введите левую границу предиктивного интервала с точностью до двух знаков после десятичной точки.

278. Зависимость спроса на некоторый вид услуг Y от его цены имеет вид $\widehat{\ln Y_i} = \underset{(2)}{30} - \underset{(0.001)}{0.03}\, P_i$. Чтобы спрос в среднем снизился на 3% цена должна увеличиться на

a) 10 единиц

b) 1 единицу

c) 10%

d) 1%

e) 100 единиц

279. Для модели $Y_i = \beta X_i + u_i$ с $\mathbb{E}(u_i) = 0$ известно, что оценка $\hat{\beta} = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i}$ обладает наименьшей дисперсией среди линейных несмещённых оценок.
Дисперсии $\mathrm{Var}(u_i)$ пропорциональны

a) $1/X_i^2$
b) $\sqrt{X_i}$
c) X_i
d) X_i^2
e) $1/X_i$

280. Оценка МНК коэффициента регрессии без свободного члена $Y_i = \beta X_i + u_i, i = 1, \ldots, n$, где $x_i = X_i - \bar{X}, y_i = Y_i - \bar{Y}$, находится по формуле

a) $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$
b) $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n X_i^2}$
c) $\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$
d) $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}$
e) $\hat{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n Y_i^2}$

281. Василий сменил единицы измерения ряда с тысяч рублей на рубли.
На какое число при этом домножится средняя абсолютная ошибка наивной модели?

282. Выберите верное утверждение про метод максимального правдоподобия.

a) Метод подбирает значения параметров, минимизирующие вероятность наибольшей ошибки прогнозов.
b) Метод подбирает наиболее вероятные значения параметров.
c) Метод подбирает значения параметров, максимизирующие вероятность наименьшей ошибки прогнозов.
d) Метод минимизирует сумму квадратов ошибок прогнозов.
e) Метод подбирает значения параметров, при которых вероятность имеющейся выборки максимальна.

283. Рассмотрим лаговые многочлены $P(L) = 1 - 0.2L$ и $Q(L) = 1 + 6L + \alpha L^2$.
При каком значении α многочлены будут сократимыми?
Ответ вводите с точностью до двух знаков после десятичной точки.

284. Элеонора исследует зависимость цены номера в отеле от звёздности отеля, $star$, (от 1 до 3 звёзд) и расстояния до моря, $dist$. Элеонора хочет оценить модель вида $price_i = \beta_1 + \beta_2 star_i + \beta_3 dist_i + u_i$.
Чтобы считаться богиней эконометрики Элеоноре стоит

a) заменить переменную $star_i$ на дамми-переменные one_i, two_i и $three_i$, равные 1 для отелей с одной, двумя и тремя звёздами соответственно
b) заменить переменную $star_i$ на дамми-переменные one_i и two_i, равные 1 для отелей с одной и двумя звёздами соответственно
c) добавить дамми-переменные one_i, two_i и $three_i$, равные 1 для отелей с одной, двумя и тремя звёздами соответственно
d) добавить в модель переменную $z_i = star_i^2$, так как эффект звёздности наверняка нелинейный
e) добавить в модель переменную $z_i = star_i \cdot dist_i$
f) использовать МНК для оценки данной модели

285. Рассмотрим модель, стоящую за тета-методом,

$$\begin{cases} y_t = \ell_{t-1} + 2 + u_t \\ \ell_t = \ell_{t-1} + 2 + 0.3u_t \end{cases}.$$

Перепишите эту модель в виде

$$\Delta y_t = \theta_1 + \theta_2 u_{t-1} + u_t.$$

В ответ укажите коэффициент θ_1.

286. Исследователь Иа-Иа использует мультипликативное разложение ряда. Очередное наблюдение равно 60, трендовая составляющая равна 40, сезонная составляющая 2.
Найдите остаточную компоненту с точностью до двух знаков после десятичной точки.

287. Петя и Вася строят доверительные интервалы для коэффициента наклона в модели парной регрессии. Оба строят интервалы по общему набору данных в 1000 наблюдений. Петя строит 95%-й интервал, а Вася — в два раза более узкий.
Какую вероятность накрытия имеет Васин интервал?
Ответ вводите с точностью до двух знаков после десятичной точки.
Можно использовать функции из статистических пакетов или таблицы.

288. Пусть u_t - белый шум. Тогда стационарным будет следующий процесс

a) $Y_t = \sum_{i=0}^{10} u_{t-i}$
b) $Y_t = Y_{t-1} - u_t$
c) $Y_t = 2018t + u_t$
d) $Y_t = tu_t$
e) $Y_t = 2Y_{t-1} - u_t$

289. Для какой цели используют LASSO?

a) Сильно увеличить дисперсию оценок за счёт небольшого смещения оценок.
b) Сильно уменьшить смещение оценок за счёт небольшого падения дисперсии оценок.
c) Сильно уменьшить смещение оценок за счёт небольшого роста дисперсии оценок.
d) Сильно уменьшить дисперсию оценок за счёт небольшого смещения оценок.
e) Сильно увеличить дисперсию оценок без смещения оценок.
f) Сильно уменьшить дисперсию оценок без смещения оценок.

290. Рассмотрим алгоритм градиентного бустинга над решающими деревьями для задачи регрессии. Какую целевую переменную «учится» прогнозировать дерево номер 4?

a) разницу между исходной переменной y_t и прогнозом третьего дерева, домноженным на темп обучения
b) исходную переменную y_t
c) нет верного ответа
d) разницу между исходной переменной y_t и прогнозом третьего дерева
e) разницу между исходной переменной y_t и суммой прогнозов первых трёх деревьев
f) разницу между исходной переменной y_t и суммой прогнозов первых трёх деревьев, домноженной на темп обучения

291. Случайная величина X имеет χ^2-распределение с шестью степенями свободы.
Найдите число a, такое что $\mathbb{P}(X > a) = 0.05$.
Можно использовать статистические функции в R/Python или других программах. Ответ вводите с точностью до двух знаков после десятичной точки.

292. В модели функция правдоподобия в точке максимума равна 10^{-10} а критерий Акаике равен 62.05. Сколько параметров оценивается в модели?

293. В модели множественной регрессии $\hat{y}_i = 2 + 3x_i + 7z_i - 6w_i$ оценённая сумма квадратов равна 777. Для оценивания использовали 200 наблюдений. Остатки обозначим вектором $\hat{u}$.
Найдите сумму $\sum_i \hat{y}_i \hat{u}_i$.

294. Выберите верный признак мультиколлинеарности.

   a) Каждый регрессор незначим по отдельности, регрессия в целом не значима.
   b) Часть регрессоров значима по отдельности, часть регрессоров — не значима по отдельности.
   c) Каждый регрессор значим по отдельности, регрессия в целом значима.
   d) Каждый регрессор значим по отдельности, регрессия в целом не значима.
   e) Каждый регрессор незначим по отдельности, однако регрессия в целом значима.

295. С помощью t-теста проверяется гипотеза о том, что

   a) стандартная ошибка коэффициента регрессии равна единице
   b) коэффициент регрессии равен единице
   c) оценка стандартной ошибки коэффициента регрессии равна единице
   d) надо ходить на семинары по эконометрике
   e) оценка коэффициента регрессии равна единице

296. Истинной является модель $Y_i = \beta_1 + \beta_2 X_i + u_i$. Глафира оценивает две регрессии: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ и $\hat{Y}_i = \hat{\gamma}_1 + \hat{\gamma}_2 X_i + \hat{\gamma}_3 Z_i$ с помощью МНК. Для коэффициента β_2

   a) оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются неэффективными
   b) оценка $\hat{\beta}_2$ является смещённой, а оценка $\hat{\gamma}_2$ — несмещённой
   c) оценка $\hat{\beta}_2$ является несмещённой, а оценка $\hat{\gamma}_2$ — смещённой
   d) оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются эффективными
   e) оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются несмещёнными

297. Взятием разностей может быть сведен к стационарному

   a) как временной ряд с детерминированным трендом, так и со случайным трендом
   b) только коинтегрированный ряд
   c) только временной ряд со случайным трендом
   d) только временной ряд с детерминированным трендом
   e) ни временной ряд с детерминированным трендом, ни со случайным трендом

298. У меня набор данных из трёх наблюдений: $x = (1, 2, 3)$, $y = (5, 6, 10)$.
Чему равна МНК-оценка $\hat{\beta}_x$ в модели $y_i = \beta_1 + \beta_x x_i + u_i$?
Ответ вводите с точностью до двух знаков после запятой.

299. Пантелеймон оценил модель парной регрессии

$$\ln \hat{y}_i = 0.4 + 1.2 \ln x_i.$$

Какая интерпретация уравнения является верной?

a) с ростом x на один процент величина y растёт в среднем на 1.2 единицы
b) с ростом x на один процент величина y растёт в среднем на 1.2 процента
c) нет верного ответа
d) с ростом x на единицу величина y растёт в среднем на 120 процентов
e) с ростом x на единицу величина y растёт в среднем на 1.2 процента
f) с ростом x на единицу величина y растёт в среднем на 1.2 единицы

300. Исследовательница Василиса построила парную регрессию $\hat{y}_i = 2 - 3x_i$ с $R^2 = 0.64$.
Найдите выборочную корреляцию между зависимой переменной y и предиктором x.

301. Гюльчатай оценила одну и ту же модель зависимости зарплаты от опыта работы по трём разным группам наблюдений.
Для 1000 сельских жителей:

$$wage_i = \beta_1 + \beta_2 exp_i + u_i, RSS_a = 150.$$

Для 2004 городских жителей:

$$wage_i = \gamma_1 + \gamma_2 exp_i + u_i, RSS_b = 200.$$

Для всех жителей сразу:

$$wage_i = \delta_1 + \delta_2 exp_i + u_i, RSS_c = 400.$$

Найдите значение F-статистики теста Чоу для проверки гипотезы об одинаковой зависимости для всех жителей.
Ответ укажите с точностью до двух знаков после десятичной точки.

302. Процесс (y_t) стационарный.
Найдите максимальное значение t, при котором гарантированно выполнено условие

$$\mathrm{Cov}(y_{108}, y_{109}) = \mathrm{Cov}(y_{120}, y_t).$$

303. Модель $y_i = \beta_1 + \beta_x x_i + u_i$ оценивают с помощью МНК. Известно, что $\bar{x} = 2$, $\bar{y} = 4$, и $\hat{\beta}_1 = 2$.
Чему равна МНК-оценка $\hat{\beta}_x$?

304. У Винни-Пуха был временной ряд из 1000 наблюдений. Винни использует алгоритм случайного леса, причем в качестве предикторов он берёт лаги y_{t-1}, ..., y_{t-9}.
По скольки наблюдениям обучается каждое из деревьев?

305. Для модели парной регрессии $y = \beta_1 s + \beta_2 x + u$, где $y = (y_1, \ldots, y_n), x = (x_1, \ldots, x_n), s = (1, \ldots, 1), \hat{u} = (\hat{u}_1, \ldots, \hat{u}_n), \hat{y} = \hat{\beta}_1 s + \hat{\beta}_2 x, \hat{u} = y - \hat{y}$ в пространстве $\mathbb{R}^n$ ортогональны вектора

a) y и $\hat{y}$
b) $\hat{u}$ и s
c) x и $\hat{y}$
d) u и $\hat{y}$
e) y и s

306. Перед Винни-Пухом временной ряд потребления мёда с восходящим трендом и сезонными колебаниями растущей амплитуды. Какая модель из предложенных лучше подходит для описания данного ряда?

a) ETS(MMN)
b) ETS(AAA)
c) ETS(MAM)
d) ETS(MNM)
e) ETS(AAN)
f) ETS(AAdA)

307. Выберите утверждение про модель локальной линейной регрессии LOESS.

a) LOESS можно оценить только для предикторов в малой окрестности нуля.
b) В качестве ядерной функции разумно выбрать любую монотонно возрастающую функцию.
c) Оценки коэффициентов бета с крышкой зависят от точки, для которой строится прогноз.
d) LOESS позволяет оценить коэффициенты, но не позволяет строить прогнозы.
e) LOESS работает гораздо быстрее, чем аналогичная парная регрессия.
f) LOESS можно использовать только на данных временных рядов.

308. Гипотеза о том, что одновременно $\beta_1 + \beta_2 = 1$ и $\beta_3 = 0$ в множественной линейной регрессии построенной по n наблюдениям проверяется с помощью статистики, имеющей распределение

a) t_{n-2}
b) t_n
c) F
d) $N(0;1)$
e) Демешева-Мамонтова
f) t_{n-k}

309. Оценки МНК вектора коэффициентов регрессии $Y = X\beta + u$ находятся по формуле

a) $(X'X)^{-1}YX$
b) $X'Y(X'X)^{-1}$
c) $(XX')^{-1}X'Y$
d) $(X'X)^{-1}X'Y$
e) $(XX')^{-1}Y'X$

310. Частным случаем какой модели является модель за тета-методом?

a) ETS(MMM) с $\beta = 0$
b) ETS(AAN) с $\beta = 0$
c) ETS(AAA) с $\gamma = 1$
d) нет верного ответа
e) ETS(AAA) с $\alpha = 1$
f) ETS(MMM) с $\alpha = 0$

311. Какая модель соответствует аддитивному методу Хольта-Винтерса?

a) ETS(ANA)
b) ETS(AAA)
c) ETS(AAdA)
d) нет верного ответа
e) ETS(AAN)
f) ETS(ANN)

312. Абу Али Яхья ибн Галиб аль-Хайят оценил множественную регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_x x_i + \hat{\beta}_z z_i$. Обозначим вектор остатков с помощью $\hat{u}$.
Выберите пары ортогональных векторов.

a) $y \perp \hat{u}$ и $y \perp y + \hat{u}$
b) $\hat{y} \perp x$ и $\hat{y} \perp x + z$
c) $x \perp y$ и $x \perp y + z$
d) $\hat{u} \perp x$ и $\hat{u} \perp x + z$
e) $\hat{u} \perp y$ и $\hat{u} \perp y + z$
f) $y \perp x$ и $y \perp x + z$

313. Рассмотрим процесс $Y_t = -0.2Y_{t-1} + u_t$. 5-ое значение автокорреляционной функции равно

a) 0.2
b) 0.00032
c) -0.2
d) 0
e) -0.00032