

Содержание

	Стр.
1. Вопрос №1	4
2. Вопрос №2	4
3. Вопрос №3	4
4. Вопрос №4	4
5. Вопрос №5	5
6. Вопрос №6	5
7. Вопрос №7	5
8. Вопрос №8	5
9. Вопрос №9	6
10.Вопрос №10	6
11.Вопрос №11	6
12.Вопрос №12	7
13.Вопрос №13	7
14.Вопрос №14	7
15.Вопрос №15	7
16.Вопрос №16	7
17.Вопрос №17	8
18.Вопрос №18	8
19.Вопрос №19	8
20.Вопрос №20	8

21.Вопрос №21	8
22.Вопрос №22	9
23.Вопрос №23	9
24.Вопрос №24	9
25.Вопрос №25	9
26.Вопрос №26	9
27.Вопрос №27	10
28.Вопрос №28	10
29.Вопрос №29	10
30.Вопрос №30	10
31.Вопрос №31	10
32.Вопрос №32	11
33.Вопрос №33	11
34.Вопрос №34	11
35.Вопрос №35	11
36.Вопрос №36	11
37.Вопрос №37	12
38.Вопрос №38	12
39.Вопрос №39	12
40.Вопрос №40	12
41.Вопрос №41	12

42.Вопрос №42	13
43.Вопрос №43	13
44.Вопрос №44	13
45.Вопрос №45	13
46.Вопрос №46	13
47.Вопрос №47	14
48.Вопрос №48	14
49.Вопрос №49	14
50.Вопрос №50	14
51.Вопрос №51	14
52.Вопрос №52	15

1 Вопрос №1

Формулировка

Формы представления данных и цели анализа данных.

Ответ

По форме представления данных можно выделить выборки, функции, семантические данные (тексты, изображения). В качестве основных целей анализа данных можно отметить анализ глобальных событий и явлений, компактное представление данных, прогнозирование.

2 Вопрос №2

Формулировка

Характеристики положения данных.

Ответ

1) Выборочное среднее, 2) $\text{med } x = x$? $x = x[k+1] \parallel x = x[k]$, в случае нечётности и чётности соответственно (выборочная медиана), 3) Z_R = полусумма экстремальных значений, 4) Полусумма выборочных квартилей $Z_Q = \frac{z_{1/4} + z_{3/4}}{2}$, 5) $Z_{tr} = \frac{1}{n-2r} \sum_{i=r+1}^{n-r} x(i)$ - усечённое среднее, 6) Среднее геометрическое $Z_G = x_1 * \dots * x_n^{1/n}$, 7) Среднее гармоническое $Z_M = \frac{1}{1/n * \sum_{i=1}^n 1/z_i}$, 8) Среднее по Колмогорову $z_k = g^{-1}(\frac{1}{n} \sum_{i=1}^n g(x_i))$.
Вариационный ряд данных - неубывающая выборка.

3 Вопрос №3

Формулировка

Характеристики рассеивания данных.

Ответ

1) Среднеквадратичное отклонение и выборочная дисперсия, 2) среднее абсолютное отклонение от медианы $d = 1/n \sum_{i=1}^n |x_i - \text{med } x|$, 3) $R = X_n - X_1$ - размах выборки, 4) интервальная ширина, 5) Медианное абсолютное отклонение

4 Вопрос №4

Формулировка

Оптимизационный подход к построению x-к положения и рассеивания данных.

Ответ

Нет ответа.

5 Вопрос №5

Формулировка

Характеристики взаимосвязи данных.

Ответ

Коэффициент корреляции Пирсона (линейная зависимость), ранговый коэффициент Спирмена - мера монотонной зависимости (не линейной), квадратный (знаковый) коэффициент корреляции.

6 Вопрос №6

Формулировка

Характеристики экстремальных значений данных.

Ответ

Нужно для выявления выбросов.

1) $|x_i - \text{med}x| > K * \text{MAD}$ $x \rightarrow \text{IQR}$

2) x_i - выброс, если $|x_i - LQ - 3/2IQR| < \min(x_n, UQ + 3/2IQR)$.

7 Вопрос №7

Формулировка

Графическое представление данных — боксплот Тьюки.

Ответ

$\text{IQR} = UQ - LQ$. x_i - выброс, если $|x_i - LQ - 3/2IQR| < \min(x_n, UQ + 3/2IQR)$.

8 Вопрос №8

Формулировка

Характеристики распределений данных: «ядерные» оценки плотности.

Ответ

Эмпирическая функция распределения, дельта функция.

Ядро - функция $K(u)$, если: 1. $K(u) \geq 0$, 2. $K(-u) = K(u)$, 3. $\int_{-\infty}^{\infty} K(u) du = 1$.

Если функция обладает первым свойством, то результатом ядерной оценки плотности действительно будет плотность вероятности. Третье свойство гарантирует, что среднее значение распределения равно среднему использованной выборки.

Нужно для оценки плотности распределения. $f(x) = \lim_{n \rightarrow \infty, h \rightarrow 0} 1/n * h \sum_{i=1}^n K(\frac{x-x_i}{h})$

9 Вопрос №9

Формулировка

Что такое точечная оценка?

Ответ

Оценка параметра — соответствующая числовая характеристика, рассчитанная по выборке. Оценки параметров генеральной совокупности делятся на два класса: точечные и интервальные.

Когда оценка определяется одним числом, она называется точечной оценкой. Точечная оценка, как функция от выборки, является случайной величиной и меняется от выборки к выборке при повторном эксперименте.

К точечным оценкам предъявляют требования, которым они должны удовлетворять, чтобы хоть в каком-то смысле быть «доброкачественными». Это несмещённость, эффективность и состоятельность.

Несмещённость - если мат ожидание оценки равно оцениваемому параметру генеральной совокупности. Эффективность - если обладает минимальной дисперсией среди всех несмещенных точечных оценок. Состоятельность - если при увеличении выборки стремится по вероятности к параметру генеральной совокупности. Генеральная совокупность - параметр от которого зависит выборка.

10 Вопрос №10

Формулировка

Что такое статистика?

Ответ

Статистика — отрасль знаний, наука, в которой излагаются общие вопросы сбора, измерения, мониторинга, анализа массовых статистических (количественных или качественных) данных и их сравнение; изучение количественной стороны массовых общественных явлений в числовой форме.

11 Вопрос №11

Формулировка

Какая оценка называется состоятельной, несмещенной, эффективной, робастной?

Ответ

Несмещённость - если мат ожидание оценки равно оцениваемому параметру генеральной совокупности. Эффективность - если обладает минимальной дисперсией среди всех несмещенных точечных оценок. Состоятельность - если при увеличении выборки стремится по вероятности к параметру генеральной совокупности. Генеральная совокупность - параметр от которого зависит выборка. Робастная - устойчивая к выбросам

12 Вопрос №12

Формулировка

Какая из двух оценок считается более эффективной?

Ответ

Для НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН ВЕРОЯТНОСТЬ ТОГО, ЧТО ТОЧЕЧНАЯ ОЦЕНКА (ширина доверительного интервала равна 0) СОВПАДЕТ С ЛЮБЫМ ЗАДАННЫМ ЗНАЧЕНИЕМ ИЛИ ОЦЕНИВАЕМЫМ ПАРАМЕТРОМ РАВНА 0.

Таким образом, точечная оценка имеет смысл лишь тогда, когда приведена характеристика рассеяния этой оценки (дисперсия). В противном случае она может служить лишь в качестве исходных данных для построения интервальной оценки. Интервальная оценка лучше.

13 Вопрос №13

Формулировка

Что такое эффективность, относительная эффективность, асимптотическая эффективность оценки?

Ответ

Нет ответа.

14 Вопрос №14

Формулировка

Что такое процедура «складного ножа»?

Ответ

Нет ответа.

15 Вопрос №15

Формулировка

Приведите примеры состоятельных оценок м.о. нормального распределения.

Ответ

Нет ответа.

16 Вопрос №16

Формулировка

Приведите примеры состоятельных оценок м.о. распределения Лапласа.

Ответ

Нет ответа.

17 Вопрос №17

Формулировка

Приведите примеры состоятельных оценок м.о. равномерного распределения.

Ответ

Нет ответа.

18 Вопрос №18

Формулировка

Приведите примеры состоятельных оценок центра симметрии распределения Коши.

Ответ

Нет ответа.

19 Вопрос №19

Формулировка

Приведите примеры состоятельных оценок стандартного отклонения нормального распределения.

Ответ

Нет ответа.

20 Вопрос №20

Формулировка

Приведите примеры состоятельных оценок дисперсии нормального распределения.

Ответ

Нет ответа.

21 Вопрос №21

Формулировка

Приведите примеры состоятельных оценок стандартного отклонения распределения Лапласа.

Ответ

Нет ответа.

22 Вопрос №22

Формулировка

Приведите примеры состоятельных оценок дисперсии распределения Лапласа.

Ответ

Нет ответа.

23 Вопрос №23

Формулировка

Приведите примеры состоятельных оценок стандартного отклонения равномерного распределения.

Ответ

Нет ответа.

24 Вопрос №24

Формулировка

Приведите примеры состоятельных оценок дисперсии равномерного распределения.

Ответ

Нет ответа.

25 Вопрос №25

Формулировка

Приведите примеры несмещенных оценок м.о. нормального распределения.

Ответ

Нет ответа.

26 Вопрос №26

Формулировка

Приведите примеры несмещенных оценок дисперсии нормального распределения.

Ответ

Нет ответа.

27 Вопрос №27

Формулировка

Назовите состоятельные оценки начальных моментов распределений.

Ответ

Нет ответа.

28 Вопрос №28

Формулировка

Назовите состоятельные оценки центральных моментов распределений.

Ответ

Нет ответа.

29 Вопрос №29

Формулировка

Назовите состоятельные оценки генеральных квантилей распределений.

Ответ

Нет ответа.

30 Вопрос №30

Формулировка

Что такое неравенство Рао-Крамера? В чем состоит его смысл?

Ответ

Нет ответа.

31 Вопрос №31

Формулировка

При каком условии достигается равенство в неравенстве Рао-Крамера? Приведите примеры.

Ответ

Нет ответа.

32 Вопрос №32

Формулировка

Сформулируйте метод максимума правдоподобия. Какова эвристическая идея этого метода?

Ответ

Нет ответа.

33 Вопрос №33

Формулировка

Сформулируйте метод моментов.

Ответ

Нет ответа.

34 Вопрос №34

Формулировка

Сформулируйте метод квантилей.

Ответ

Нет ответа.

35 Вопрос №35

Формулировка

Каковы общие свойства оценок максимума правдоподобия?

Ответ

Нет ответа.

36 Вопрос №36

Формулировка

Каковы оценки максимума правдоподобия параметров нормального распределения?

Ответ

Нет ответа.

37 Вопрос №37

Формулировка

Каковы оценки максимума правдоподобия параметров равномерного распределения?

Ответ

Нет ответа.

38 Вопрос №38

Формулировка

Какова оценка максимума правдоподобия вероятности «успеха» биномиального распределения?

Ответ

Нет ответа.

39 Вопрос №39

Формулировка

Какова оценка максимума правдоподобия для параметра масштаба показательного распределения?

Ответ

Нет ответа.

40 Вопрос №40

Формулировка

Каковы оценки максимума правдоподобия параметров распределения Лапласа?

Ответ

Нет ответа.

41 Вопрос №41

Формулировка

Каковы оценки метода моментов параметров нормального распределения?

Ответ

Нет ответа.

42 Вопрос №42

Формулировка

Что такое доверительный интервал?

Ответ

Нет ответа.

43 Вопрос №43

Формулировка

Что такое интервальная оценка параметра и каковы ее отличия от точечной оценки?

Ответ

Нет ответа.

44 Вопрос №44

Формулировка

Что такое точность и надежность интервальной оценки?

Ответ

Нет ответа.

45 Вопрос №45

Формулировка

Что такое критерий согласия?

Ответ

Нет ответа.

46 Вопрос №46

Формулировка

Что такое ошибки первого и второго рода?

Ответ

Нет ответа.

47 Вопрос №47

Формулировка

Какие критерии согласия вы знаете?

Ответ

Нет ответа.

48 Вопрос №48

Формулировка

Какова общая схема проверки статистических гипотез с использованием критериев согласия?

Ответ

Нет ответа.

49 Вопрос №49

Формулировка

Каково происхождение термина «регрессия»?

Ответ

Нет ответа.

50 Вопрос №50

Формулировка

Что такое задача простой линейной регрессии?

Ответ

Нет ответа.

51 Вопрос №51

Формулировка

Какие методы оценивания параметров простой линейной регрессии вы знаете?

Ответ

Нет ответа.

52 Вопрос №52

Формулировка

Как проверяется адекватность полученного решения задачи простой линейной регрессии?

Ответ

Нет ответа.