1. Формула Стерджеса: - Зачем используется формула Стерджеса при анализе статистических данных?

*Ответ: Для определения интервалов классов при составлении гистограммы.*

1. Модель Lasso позволяет …

*Ответ: обнулить малозначимые коэффициенты регрессии*

1. Эксцесс и асимметрия: - Если распределение левостороннее, то значение асимметрии будет:

*Ответ: Меньше 0*

1. Коэффициент регуляризации λ2  с ненулевым значением используется в моделях линейной регрессии …

*Ответ: Ridge*

1. Что называется статистической гипотезой?

*Ответ: Предположение о параметрах распределения случайной величины*

1. Случайная величина, которая может принимать любое значение из непрерывного множества, называется:

*Ответ: Непрерывной*

1. Что называется значением p-value?

*Ответ: Вероятность попадания статистики критерия в критическую область при условии, что основная гипотеза верна*

1. Коэффициент детерминации описывает …

*Ответ: точность модели*

1. Укажите функцию Excel, используемую для расчёта критической точки при проверке гипотезы о виде распределения с помощью критерия Пирсона:

*Ответ: ХИ2ОБР()*

1. В какой из предметных областей статистика НЕ играет ключевой роль?

*Ответ:* Искусство

1. Рассчитайте выборочную дисперсию (смещенную) по выборке:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 8 | 7 | 9 | 13 | 8 |

Ответ указать с точностью до двух знаков! Разделитель - точка. Примеры: 3.21, 2.00, 1.50

*Ответ: 4.4*

1. При проверке гипотезы H0: m=5, при H1: m<5, при известной дисперсии, равной 1, было получено значение статистики критерия 10 и критическая точка 3. Укажите результат проверки данной гипотезы.

*Ответ: Основная гипотеза принимается, т.к. значение статистики критерия не попало в критическую область*

1. Дана частотная таблица сопряженности для случайных величин X,Y. Рассчитайте статистику Хи-квадрат (значение критерия Пирсона) для данной таблицы.

**x\y 1 2**

* + 1. 3 2
    2. 1 4

Ответ на вопрос с расчетом указывается с точностью до двух знаков после разделителя , разделитель - точка, используются стандартные правила округления (0.567 = 0.57, 0.877 = 0.88 и т.п.). Если получается целое число, то можно вписывать его как с разделителем, так и без. Например: -1.00, 0.56, 0.50, 0.33, 2

*Ответ: 5/3=1.67*

1. Ожидается, что добавление специальных веществ уменьшит жесткость воды. По оценке жесткости воды до и после добавления специальных веществ по 40-ка и 50-ти пробам соответственно получим средние значения жесткости (в стандартных единицах), равные 4,0 и 0,8. Дисперсия измерений в обоих случаях предполагается равной 0,25. Принять α=0,05. Контролируемая величина имеет нормальное распределение. Найти разность критического значения и наблюдаемого значения.

*Ответ: -28.21*

1. Из 200 задач первого раздела курса математики, предложенных для решения, абитуриенты решили 130, а из 300 задач второго раздела абитуриенты решили 120. Можно ли при α=0,01 утверждать, что первый раздел школьного курса абитуриенты усвоили лучше, чем второй? В ответ записать разность критического значения и наблюдаемого значения.

*Ответ: -3.15*

1. Основная и альтернативная гипотеза. Выбрать правильные утверждения:

*Ответ:* *Статистическая гипотеза называется основной, если она проверяема; Альтернативная гипотеза – это отрицание нулевой гипотезы*

1. Если два набора данных имеют одинаковый межквартильный размах, какие из следующих характеристик также имеют оба этих набора?

*Ответ: Расстояния от первого к третьему квартилям в обоих наборах будут одинаковыми*

1. Если X — случайная величина с ожиданием E(X) = μ и стандартным отклонением σ, то E(2X + 3) равно:

*Ответ: E(2X+3)=2E(X)+3=2μ+3*

1. Для моделирования равномерно распределённой случайной величины используют:

*Ответ: метод обратной функции*

1. Функция Excel ДИСП(var), если диапазон var распределён нормально с параметрами m = 1, sigma = 2 даст результат, близкий к:

*Ответ: 4*

1. Математическая статистика решает следующие задачи:
   * + 1. систематизация полученного статистического материала (этап описания массового явления);
       2. выявление свойств и закономерностей изучаемого явления (этап анализа и прогноза).

*Ответ: Оба ответа верны*

1. Критерий согласия Пирсона основан на:

*Ответ: подсчёте статистики, распределённой по закону Хи-квадрат, если основная гипотеза справедлива*

1. Гипотеза о том, что математическое ожидание некоторой нормально распределенной случайной величины Mx > 20, является простой.

*Ответ: Неверно*

1. Если в реакторе было зафиксировано 20 случаев выброса радиации за год, какое распределение лучше всего подойдёт для моделирования этих событий?

*Ответ: Биномиальное*

1. Рассчитайте границы доверительного интервала для оценки генерального среднего, при уровне значимости alpha = 1% и известной генеральной дисперсии D = 4 по выборке из нормально распределенной ГС:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1,2 | 2,4 | 0,86 | -0,52 | 2,49 |

Введите нижнюю границу или верхнюю (только одну, по выбору!) с точностью до двух знаков, разделитель - точка. Примеры: 1.00, -2.58, 3.60.

*Ответ: 3.59*

1. Фирма рассылает рекламные каталоги возможным заказчикам. Как показал опыт, вероятность того, что организация получившая каталог, закажет рекламируемое изделие, равна 0,08. Фирма разослала 1000 каталогов новой, улучшенной, формы и получила 100 заказов. Найти разность критического значения и наблюдаемого значения при уровне значимости 0,05.

*Ответ: -0.685*

1. Рассчитайте выборочную медиану по выборке:

7 10 7 12 4

Ответ указать с точностью до двух знаков! Разделитель - точка. Примеры: 3.21, 2.00, 1.50

*Ответ: 7*

1. По выборке Х объема n=30 найден средний вес m\*(X) =130 г изделий, изготовленных на первом станке; по выборке Y объема m=40 найден средний вес m\*(Y) =125 г изделий изготовленных на втором станке, причем случайные величины X и Y распределены нормально. Генеральные дисперсии этих величин известны: D\*(X)=60 г2 , D\*(Y)=80 г2 . Требуется при уровне значимости 0,01 проверить нулевую гипотезу Н0: M(X)=M(Y) при конкурирующей гипотезе Н1: M(X)≠M(Y) и найти разность критического значения и наблюдаемого значения.

*Ответ: Гипотеза принята; 0.791*

1. Все оценки, рассчитанные по методу максимального правдоподобия, являются по крайней мере…

*Ответ: состоятельными, асимптотически несмещёнными и асимптотически эффективными*

1. Рассчитайте границы доверительного интервала для оценки генерального среднего, при уровне значимости alpha = 1% и известной генеральной дисперсии D = 4 по выборке из нормально распределенной ГС:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3,9 | 2,74 | 1,21 | 0,93 | 1,00 |

Введите нижнюю границу или верхнюю (только одну, по выбору!) с точностью до двух знаков, разделитель - точка. Примеры: 1.00, -2.58, 3.60.

*Ответ: 4.26*

1. Критическая область при проверке гипотезы критерием Колмогорова для выборки из 100 элементов при alpha = 10% определяется как:

*Ответ: (0.122; ∞)*

1. Можно ли вывести на графике тренда Excel коэффициент детерминации уравнения регрессии?

*Ответ: да*

1. Что является предпосылкой для применения линейной регрессии?

*Ответ: Все перечисленное (отсутствие мультиколлинеарности, гомоскедастичность ошибок, нормальность распределения переменных)*

1. Рассчитайте границы доверительного интервала для оценки генерального среднего, при уровне значимости alpha = 1% и известной генеральной дисперсии D = 4 по выборке из нормально распределенной ГС:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -0.89 | -0.5 | 2.46 | 1.69 | 1.17 |

Введите нижнюю границу или верхнюю (только одну, по выбору!) с точностью до двух знаков, разделитель - точка. Примеры: 1.00, -2.58, 3.60.

*Ответ: 3.09*

1. Какие из следующих чисел указывают на распределение или разброс наблюдений в наборе данных?

*Ответ: Меры изменчивости*

1. Проверьте гипотезу о нормальности выборки с помощью критерия Колмогорова, если значение критерия составляет 4.5, а значение критической точки 2.03:

*Ответ: Выборка не является нормальной, т.к. значение попадает в критическую область*

1. Вы рассматриваете два ядерных реактора: А и В. Вы хотите узнать, какой из них более эффективен. Какой статистический тест вы бы использовали для сравнения их эффективности?

*Ответ: ANOVA*

1. Свойство точечной оценки, если она имеет минимально возможную дисперсию среди всех оценок, называется…

*Ответ: эффективность*

1. Какой статистический метод наиболее подходит для прогнозирования времени, через которое ядерный реактор достигнет критической температуры на основе исторических данных?

*Ответ: линейная регрессия*

1. Известно, что генеральная совокупность распределена нормально со средним квадратическим отклонением 5. Найти доверительный интервал для оценки математического ожидания а с надежностью 0,95, если выборочное среднее 24.15, а объем выборки 100.

*Ответ: 23.33<а<24.97*

1. При проверке гипотезы Н0: m=5, при Н1: m<5, при известной дисперсии, равной 1, для заданного уровня значимости 5%, было получено значение p-value = 10%. Укажите результаты проверки гипотезы.

*Ответ: гипотеза принимается, так как p-value>alpha*

1. Рассчитайте выборочную дисперсию (смещенную) по выборке:

5 6 7 10 11

*Ответ: 5.36*

1. Что называется критической областью статистики критерия при проверке статистической гипотезой?

*Ответ: Область значений статистики критерия, в которой должна приниматься альтернативная гипотеза; Область маловероятных значений статистики критерия при условии истинности основной гипотезы.*

1. Функция Excel ЭКСЦЕСС(var), если диапазон var распределён нормально с параметрами m = 1, sigma = 2 даст результат, близкий к:

*Ответ: 0*

1. Проверьте гипотезу о нормальности выборки с помощью критерия Пирсона, если значение критерия составляет 13.5, а значение критической точки 14.03

*Ответ: Выборка является нормальной, т.к. значение не попадает в критическую область*

1. При проверке гипотезы Н0: m=5, при Н1: m<5, при известной дисперсии, равной 1, для заданного уровня значимости 10%, было получено значение p-value = 12%. Укажите результаты проверки гипотезы.

*Ответ: гипотеза принимается, так как p-value>alpha*

1. Функция Excel ЧАСТОТА(var), где var – имя массива значений случайной величины, подсчитает:

*Ответ: выдаст ошибку*

1. Нормальную случайную величину моделируют на основе:

*Ответ: центральной предельной теоремы*

1. При проверке гипотезы H0: m=5, при H1: m<5, при известной дисперсии, равной 1, было получено значение статистики критерия -10 и критическая точка -3. Укажите результат проверки данной гипотезы.

*Ответ: Основная гипотеза отвергается, т.к. значение статистики критерия попало в критическую область*

1. Укажите результат применения функции Excel ХИ2ОБР(5%;0):

*Ответ: ошибка*

1. Критическая область бывает:

*Ответ: левосторонняя, правосторонняя, двусторонняя*

1. При проверке гипотезы Н0: m=5, при Н1: m<5, при известной дисперсии, равной 1, для заданного уровня значимости 1%, было получено значение p-value = 0,1%. Укажите результаты проверки гипотезы.

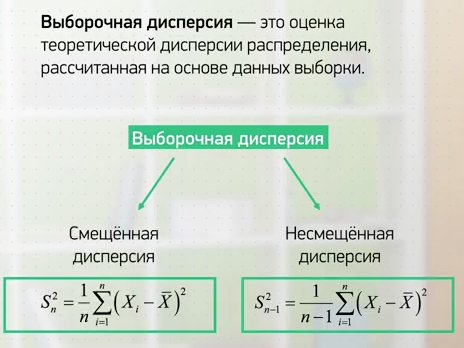
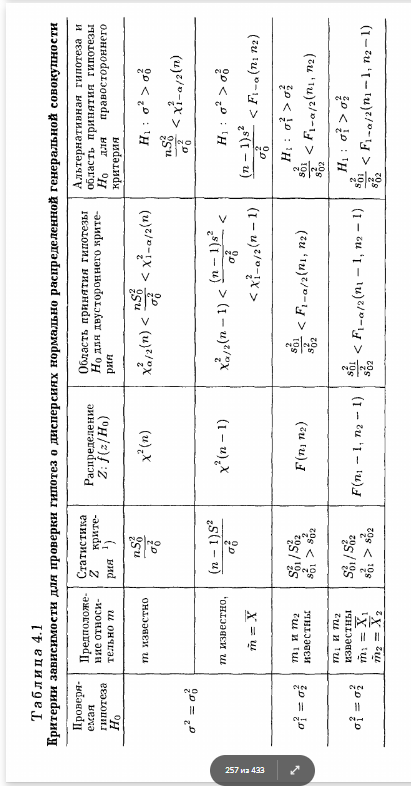
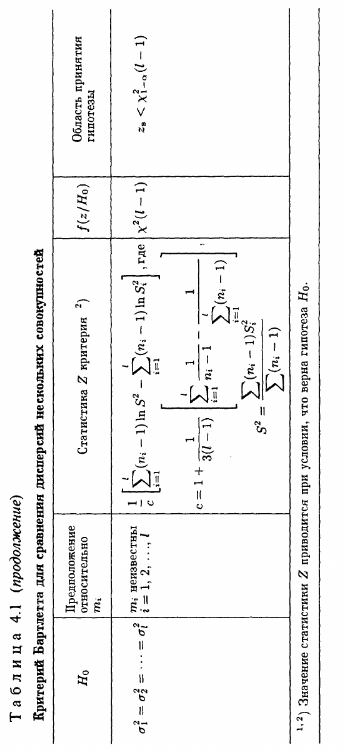
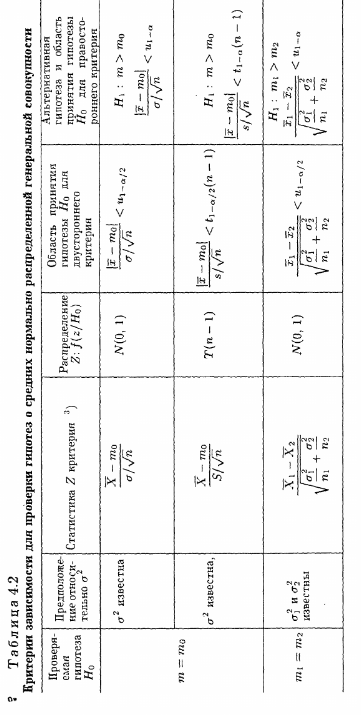
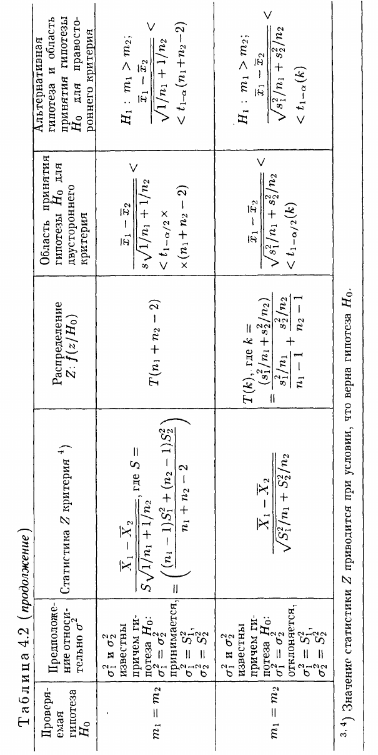
*Ответ: гипотеза не принимается, так как p-value<alpha*

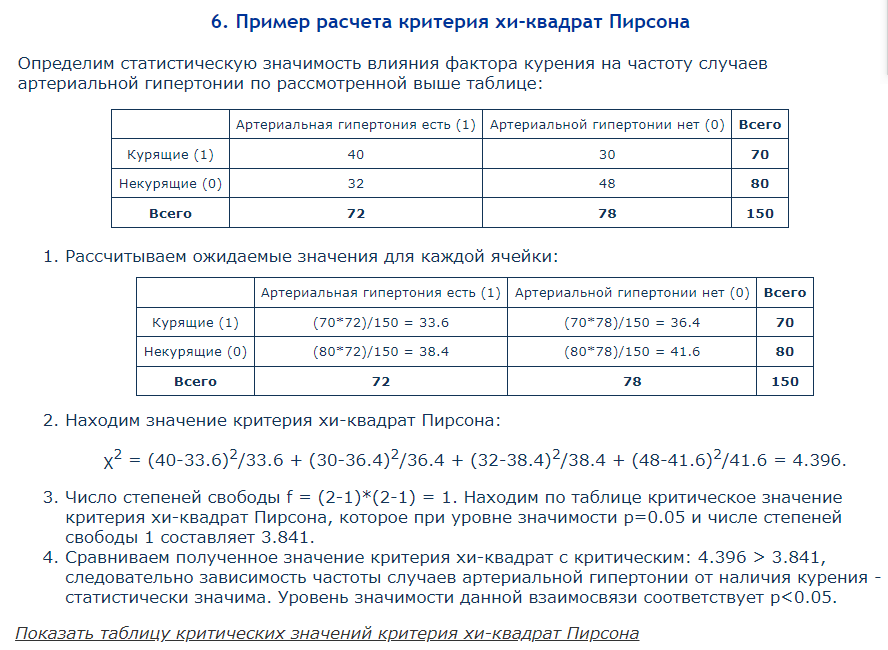
1. Рассчитайте коэффициент корреляции Пирсона для выборок Х и Y, если известны: среднее Х = -0.05, среднее У = -0,09, среднее ХУ = 0,04, ско Х = 1,01, ско У = 0,88.

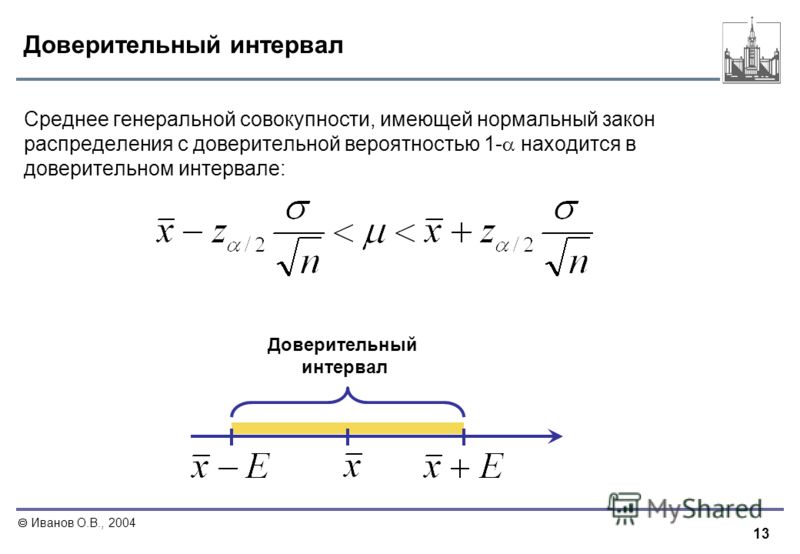
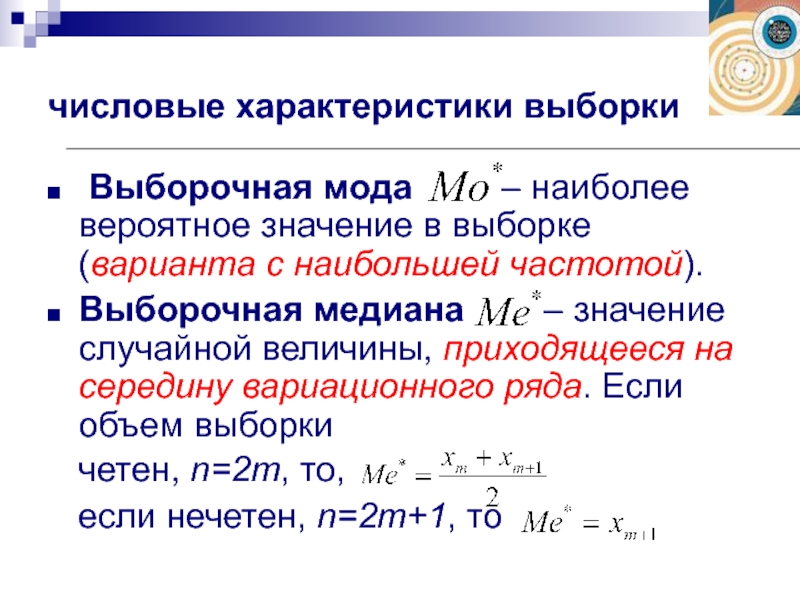
Введите нижнюю границу или верхнюю (только одну, по выбору!) с точностью до двух знаков, разделитель - точка. Примеры: 1.00, -2.58, 3.60.

*Ответ: 0.04*

Полезные ссылки:

* <https://allcalc.ru/node/1835> - несмещенная выборочная дисперсия, калькулятор
* 
* S02- смещенная дисперсия
* Проверка гипотезы
* Пример расчета критерия Пирсона



* <https://www.matburo.ru/Examples/Files/ms_pg_10.pdf> - задача про воду
* <https://www.matburo.ru/Examples/Files/ms_pg_2.pdf> - задача про абитуриентов
* Доверительный интервал среднего
* 
* <https://www.matburo.ru/Examples/Files/ms_pg_8.pdf> - задача о рекламе в фирмах
* Выборочная медиана и мода
* <https://moodle.kstu.ru/pluginfile.php/514891/mod_resource/content/6/Лекция.%20%20Статистика.%20Проверка%20статистических%20гипотез.%20Часть%202.pdf> – задача про средний вес изделий
* <https://datalearning.ru/index.php/textbook?cid=1&mid=1&topic=1> – характеристики оценок
* <https://koi.tspu.ru/biostat/Kolmogorov-Smirnov%20statistics.pdf> – критические значение Колмогорова (они должны быть больше, чем известное нам для попадания в область)
* <https://cyberleninka.ru/article/n/osnovnye-predposylki-klassicheskoy-lineynoy-regressii-i-posledstviya-ih-narusheniy/viewer> - предпосылки линейной регрессии
* Коэффициент корреляции Пирсона
* 