**Задания университетской олимпиады по анализу данных - 2016**

**I тур**

1. Распределение дискретного случайного вектора  с независимыми компонентами  и  задано таблицей распределения

|  |  |
| --- | --- |
|  |  |
| 0 | 1 | 2 |
| -1 | 0,40 | 0,20 | a |
| 1 | b | 0,05 | c |

Найдите a, b, c.

1. Когда математик Александр отправляется по воскресеньям в магазин, жена дает ему список с продуктами. Известно, что функция распределения случайной величины  – общей стоимости продуктов в списке – имеет вид



а) Сколько денег должен взять Александр, чтобы с вероятностью 0,9 он смог оплатить покупку?

б) Каким будет ответ в пункте (а), если известно, что покупка обойдется Александру не менее чем в тысячу рублей?

1. Имеется случайная выборка , все элементы  которой независимы друг с другом, распределены так же, как генеральная совокупность 

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1 | 3 | 5 |
|  | a | 0,2 | 0,8-a |

а) Каковы допустимые значения параметра a?

б) При каком значении m оценка  для параметра a будет несмещенной?

1. Для параметра  имеются две независимые несмещенные оценки  и  с дисперсиями  и  соответственно.

а) Найдите значение , при котором оценка  будет иметь наименьшую дисперсию.

б) Будет ли полученная оценка более эффективной, чем оценка  или ?

1. Математическое ожидание случайной величины  равно , дисперсия . На основе случайной выборки , элементы которой независимы и распределены так же, как генеральная совокупность, для параметра  предлагаются две оценки:





а) При каких значениях  и  обе оценки для  будут несмещенными?

б) Выберите из оценок более эффективную.

1. Статистику Кристине надоели обычные, симметричные доверительные интервалы. Она хочет построить 90% доверительный интервал для математического ожидания так, чтобы математическое ожидание недооценивалось в двух случаях из ста, а переоценивалось в восьми случаях из ста. Помогите Кристине вывести выражение для такого интервала в случае нормальной генеральной совокупности и известной дисперсии.
2. Рассматривается модель



где  – реализация случайной величины , причем  взаимно независимы и распределены нормально с параметрами .

Для оценки параметров модели  имеются выборочные данные вида

,

где  - реализация случайной величины ,

.

Будет ли прямая, проходящая через точки  и , где , , состоятельной, несмещенной и эффективной оценкой регрессии  на ?

1. С целью прогноза доли тех выпускников, которые находят работу по специальности в течение месяца после окончания вуза, в течение ряда лет был проведен анализ, на основе которого получено, что
2. среди выпускников, нашедших работы по специальности в течение месяца, средний балл в дипломе составляет 4,6, а средний опыт работы по специальности 6,1 месяца;
3. среди выпускников, не нашедших работы по специальности в течение месяца, средний балл в дипломе составляет 3,9, а средний опыт работы по специальности 1,5 месяца;
4. доля нашедших в течение месяца работу по специальности равна 0,3;
5. оценки ковариационной матрицы признаков «балл» и «опыт работы», полученные по выборкам выпускников, нашедших и не нашедших работу по специальности в течение месяца, равны между собой и равны .

 В предположении нормальности двумерного случайного вектора внутри каждого классов оценить долю тех выпускников, которые при среднем балле 4 и опыте работы по специальности 10 месяцев потратят на поиски работы по специальности не более месяца.

1. В комедии Н. В. Гоголя «Женитьба» купеческая дочь Агафья Тихоновна никак не могла выбрать из четырех женихов. Предположим, что внимательная сваха в течение часа смотрин отмечала количество благосклонных взглядов Агафьи Тихоновны на каждого из женихов:

- на Никанора Ивановича 28 взглядов;

- на Ивана Кузьмича 10 взглядов;

- на Ивана Павловича 16 взглядов;

- на Балтазара Балтазарыча 10 взглядов.

Можно ли утверждать, что кому-либо из женихов Агафья Тихоновна отдает предпочтение (на уровне значимости 0,05)?

1. Рассмотрим статистику падений самолетов-снарядов в южной части Лондона во время Второй мировой войны. Вся территория разделена на n=576 малых участков площадью 0,25 км2 каждый. Общее число снарядов составило 537. В таблице ниже представлены данные о количестве участков Nk, на которые приходилось по k падений снарядов:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 и больше |
|  | 229 | 211 | 93 | 35 | 7 | 1 |

Большинство населения Лондона верило в тенденцию точек падения к группировке (неслучайность падения снарядов). Насколько обоснованным было такое мнение?