ВИШ. Теория вероятностей и математическая статистика 2024/2025

**Midterm exam, 27.03.2025**

Пожалуйста, напишите свое имя, фамилию и номер группы на каждом листе, который содержит формулировки задач. В задачах 1- 4 вы должны написать **только ответы**. Вы должны дать ответ в виде числа (а не формулы) если вам явно не указано написать формулу. В задаче 5 вы должны написать **полное решение.**

**Часть 1**(40% от общей оценки, время: 45 минут)

**Задача 1**

(а) Дана выборка с измерениями IQ у 20 сомов.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| IQ Scores | 90 | 95 | 100 | 105 | 110 |
| 90 | 95 | 100 | 105 | 110 |
| 90 | 95 | 100 | 105 | 110 |
| 90 | 95 | 100 | 105 | 110 |

Посчитайте среднее значение, медиану, дисперсию. Ответы запишите через запятую.

Ответ:

(b) По данным из задания (a) изобразите график boxplot.

Ответ:

(с) Определите квантиль, который соответствует 72 перцентилю,

Ответ:

**Задача 2**

(a) Дана выборка ​, где каждая случайная величина подчиняется распределению Пуассона с математическим ожиданием E[X]=10. Определите функцию плотности вероятности (pdf) первой порядковой статистики .

Ответ:

(b) Из пункта (a) вычислите математическое ожидание и дисперсию.

Ответ:

(с) Пусть случайные величины X и Y описывают характеристики двух станков на заводе. Совместная плотность вероятности этих величин задана функцией:

Определите значение c , E[X] и E[Y].

Ответ:

**Задача 3**

(a) Руководство компании хочет изучить, как возможное сокращение рабочего времени (без уменьшения зарплаты) может повлиять на уровень удовлетворенности сотрудников. Для этого аналитики предложили случайным образом выбрать 600 сотрудников и разделить их на 3 группы: Сотрудники первой группы работают 4 дня в неделю вместо 5. Сотрудники второй группы работают 6 часов в день вместо 8. Сотрудники третьей группы работают в обычном режиме. На презентации проекта один из руководителей заметил, что должность сотрудников может значительно повлиять на результаты эксперимента и предложил более тщательно распределять участников по группам.

Используя только слова из описания эксперимента, укажите:

Ответ: experimental units:

independent variable:

response variable:

confounded variable:

(b) Пусть  - две несмещенные оценки параметра , . Коэффициент корреляции между оценками равен ρ(T1, T2) = 0,25.Найдите среднеквадратичную ошибку (MSE) оценки ?

Ответ:

(с) Пусть — случайная выборка из генеральной совокупности, где каждый ​ — это **время реакции** водителей (в секундах) при внезапном появлении препятствия. Предположим, что время реакции распределено равномерно: , где a и b — неизвестные параметры.

Докажите, что данная оценка является несмещенной.

Ответ:

**Задача 4**

Пусть - оценка параметра . Докажите формулу разложения среднеквадратичной ошибки MSE() на смещение и дисперсию. При доказательстве вы должны дать определение среднеквадратичной ошибки, сформулировать формулу разложения и показать, что она справедлива.

**Напишите свое решение здесь.**

**Часть 2** (60% от общей оценки, время: 45 минут)

**Задача 5**

Пусть ​ — случайная выборка из генеральной совокупности, где каждый — это индекс массы тела (BMI) случайного человека. Известно, что BMI в популяции распределён нормально:

где — истинное среднее значение BMI, а — истинная дисперсия. Рассмотрим три различные точечные оценки для параметра .

|  |  |  |
| --- | --- | --- |
| Выборочное среднее | Взвешенное среднее с учётом крайних значений | Оценка усечённого среднего |
|  |  |  |

(a) Определить, какие из этих оценок **несмещённые**.

(b) Определить, какая из них **эффективная** (имеет наименьшую дисперсию среди несмещённых оценок).

(c) Доказать, что если оценка ​ является **состоятельной**, то ​ по вероятности при .

(d) Рассмотреть случай, когда ​ является **асимптотически несмещённой**, и объяснить, чем это отличается от строгой несмещённости.

**Задача 6**

В супермаркете работают три кассы. Время обслуживания одного покупателя каждым из кассиров подчиняется **показательному распределению** с плотностью:

При этом:

Первый кассир **самый опытный**, его среднее время обслуживания **в два раза меньше**, чем у остальных двух. Это означает, что его интенсивность обслуживания ​ **в два раза выше**, чем у остальных. Второй и третий кассиры обслуживают покупателей с одинаковой интенсивностью ​. Экспериментально было зафиксировано, что: Первый кассир обслужил покупателя за **1 минуту**. Второй кассир обслужил покупателя за **2 минуты**. Третий кассир обслужил покупателя за **1.5 минуты**.

Определите:

**(а) Методом максимального правдоподобия (ММП) оценить параметр λ​** — интенсивность обслуживания кассиров.

(b) Проверьте найденную оценку на несмещенность.

(с) Проверить эффективность оценки по границе Рао-Крамера.

(d) Проверить состоятельность оценки.