Министерство образования и науки Российской Федерации

Федеральное государственное автономное образовательное учреждение

высшего образования

«Национальный исследовательский ядерный университет «МИФИ»

ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ

КАФЕДРА КИБЕРНЕТИКИ

«УтверждЕН»

на заседании кафедры 22

«\_\_\_»\_\_\_\_\_\_\_\_\_\_20\_\_ г.,

протокол №\_\_\_\_\_\_\_\_\_

.зав.каф.22\_\_\_\_\_\_\_\_\_\_\_\_\_/А.М.Загребаев/

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ**

**Математическая статистика**

|  |  |
| --- | --- |
| Направление подготовки (специальность) | 09.03.04 Программная инженерия |
|  |  |
| Профиль подготовки (при его наличии) | Математическое и программное обеспечение вычислительных машин и компьютерных сетей, |
|  |  |
| Наименование образовательной программы (специализация) | Математическое и программное обеспечение вычислительных машин и компьютерных сетей, |
|  |  |
| Квалификация (степень) выпускника | бакалавр |
|  |  |
| Форма обучения | очная |

г. Москва, 2019 г.

**ПАСПОРТ**

**фонда оценочных средств**

**по дисциплине «Математическая статистика»**

(наименование дисциплины)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Семестр** | **Трудоем-кость., кред.** | **Общий объем курса, час.** | **Лекции, час.** | **Практич. занятия, час.** | **Лаборат. работы, час.** | **СРС, час.** | **Форма(ы) итог. контроля, экз./зач./ КР/КП** |
| 5 | 6 | 216 | 32 | 16 | 16 | 98 | экзамен |

**Модели контролируемых компетенций**

В результате освоения дисциплины у выпускника формируются следующие компетенции:

| **Код компетенции** | **Компетенция** |
| --- | --- |
| ПК-12 | способность к формализации в своей предметной области с учетом ограничений используемых методов исследования |
| ПК-16 | Способность формализовать предметную область программного проекта и разработать спецификации для компонентов программного продукта |

**Программа оценивания контролируемых компетенций**

Формирование у студентов компетенций контролируется в течение всего времени освоения дисциплины в рамках:

* текущего контроля;
* рубежного контроля;
* промежуточного контроля.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ п.п.** | **Наименование раздела учебной дисциплины** | **Неде-ли** | **Лек-ции, час.** | **Практ. зан./ семи-нары, час.** | **Лаб. рабо-ты, час.** | **Обязат. текущий контроль (форма\*, неделя)** | **Аттеста-ция раздела (форма\*, неделя)** | **Макси-мальный балл за раздел \*\*** | **Компетенции по разделам, проверяемые при текущем и рубежном контроле** | **Компетенции, проверяемые на зач. /экз.** |
| 5 семестр | | | | | | | | | | |
| 1 | Проверка статистических гипотез | 1-4 | 8 | 4 | 8 |  | КИ6 | 20 | ПК-12,  ПК-16 |  |
| 2 | Анализ статистических взаимосвязей | 5-14 | 20 | 10 | 6 | КР6,ДЗ14 | КИ14 | 20 | ПК-12,  ПК-16 |  |
| 3 | Основы многомерного статистического анализа | 15-16 | 4 | 2 | 2 |  | КИ17 | 10 | ПК-12,  ПК-16 |  |
|  | Экзамен |  |  |  |  |  | Э | 50 |  | ПК-12,  ПК-16 |
|  | Итого за 5 семестр |  |  |  |  |  |  | 100 |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Недели** | **Содержание / Темы занятий** | **Лек., час.** | **Пр./сем., час.** | **Лаб., час.** | **Компетенции по темам, проверяемые при текущем контроле** | **Виды тек.контроля по проверке компетенций** | **Компетенции по темам, проверяемые на зач. /экз.** |
| 5 семестр | | | | | | | |
| 1-4 | **1. Проверка статистических гипотез.** | 8 | 4 | 8 |  |  |  |
| Проверка статистических гипотез. Терминология и примеры. Выборочные аналоги характеристик генеральной совокупности. Статистика критерия. Уровень значимости. Критическая область. Проверка гипотез о виде распределения. Критерий согласия Колмогорова. Проверка гипотез о виде распределения с помощью критерия хи-квадрат Пирсона. Определение параметров распределения. Метод обратной функции для визуальной проверки предположения о нормальности распределения.  Проверка параметрических гипотез. Статистика критерия. Критическая область. Ошибки 1-го и 2-го рода. Функции мощности.  Анализ нормальных выборок с помощью критерия отношения правдоподобия. Проверка гипотез о дисперсиях для выборок из нормальных генеральных совокупностей. Проверка гипотез о средних для выборок из нормальных генеральных совокупностей. Построение доверительных интервалов и проверка параметрических гипотез для нормальных выборок. Проверка гипотезы о значении коэффициента корреляции в двух нормальных выборках. Проверка гипотезы о вероятности в схеме Бернулли. | ПК-12,  ПК-16 | КР6, ДЗ14 | ПК-12,  ПК-16 |
| 5 | **2. Таблицы сопряженности.** | 2 | 1 | 0 |  |  |  |
| Виды связей между величинами. Функциональная и статистическая связи. Корреляционная связь. Анализ статистической связи между номинальными величинами. Эмпирическая и теоретическая таблицы сопряжённости. Проверка гипотезы о наличии статистической связи. | ПК-12,  ПК-16 | ДЗ14 | ПК-12,  ПК-16 |
| 6 | **3. Однофакторный дисперсионный анализ.** | 2 | 1 | 2 |  |  |  |
| Виды дисперсии в совокупности, разделённой на части. Эмпирический коэффициент детерминации и эмпирическое корреляционное отношение. Гипотеза дисперсионного анализа. Статистика Фишера. | ПК-12,  ПК-16 | ДЗ14 | ПК-12,  ПК-16 |
| 7-10. | **4. Корреляционный анализ.** | 8 | 4 | 2 |  |  |  |
| Статистическая связь между компонентами нормально распределенного случайного вектора. Корреляционное отношение. Оценивание коэффициента корреляции по выборочным данным. Оценивание коэффициента детерминации и корреляционного отношения по выборочным данным. Ранговые коэффициенты корреляции по Спирмену и по Кендаллу. | ПК-12,  ПК-16 | ДЗ14 | ПК-12,  ПК-16 |
| 1-14. | **5. Регрессионный анализ** | 8 | 4 | 2 |  |  |  |
| Статистические модели. Задачи регрессионного анализа. Оценивание параметров уравнения регрессии. Метод наименьших квадратов. Простейшая линейная регрессионная модель. Линейная регрессионная модель общего вида. | ПК-12,  ПК-16 | ДЗ14 | ПК-12,  ПК-16 |
| 15-16 | **6. Основы многомерного статистического анализа.** | 4 | 2 | 2 |  |  |  |
| Анализ множественный статистических взаимосвязей. Множественная линейная регрессия. Множественный коэффициент корреляции. Основы кластерного анализа многомерных данных. | ПК-12,  ПК-16 | Лр15, Сем16 | ПК-12,  ПК-16 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *5 семестр*  *Семинарские занятия* | **Компетенции по темам, проверяемые при текущем контроле** | **Виды тек.контроля по проверке компетенций** | **Компетенции по темам, проверяемые на зач. /экз.** |
| 1-4 | **Проверка статистических гипотез** | | | |
| 1. Проверка статистических гипотез. | ПК-12,  ПК-16 | Сем | ПК-12,  ПК-16 |
| 5-14 | **Анализ статистических взаимосвязей** | | | |
| 2. Анализ статистических взаимосвязей. | ПК-12,  ПК-16 | Сем | ПК-12,  ПК-16 |
| 15-16 | **Основы многомерного статистического анализа** | | | |
| 3. Основы многомерного статистического анализа. | ПК-12,  ПК-16 | Сем | ПК-12,  ПК-16 |

**Соответствие оценочных средств видам контроля**

|  |  |
| --- | --- |
| **Вид контроля** | **Наименование оценочного средства (способ оценки: устно/ письменно /комп. технолог.)** |
| ЛР | Отчет по лабораторной работе (в электронном виде) |
| Защита лабораторной работы (устно) |
| КР | Контрольная работа (письменно) |
| Сем | Решение задач (у доски, устно) |
| КИ | Контроль по итогам выполнения (интегральная оценка без проведения дополнительного контроля) |
| Э | Вопросы к экзамену, задачи к экзамену (комп. технолог.) |

|  |
| --- |
| Федеральное государственное автономное образовательное учреждение  высшего образования |
| **«Национальный исследовательский ядерный университет «МИФИ»** |
| ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ |
| КАФЕДРА КИБЕРНЕТИКИ |

**Комплект заданий для контрольной работы по дисциплине**

**«Математическая статистика»**

|  |  |  |
| --- | --- | --- |
| Составитель | Трофимов А.Г. | Доцент, к.т.н. |
| Учебный год | 2018/2019 |  |

**Задача №1**. Анализируется расход электроэнергии на сборку и наладку 10 медицинских ультразвуковых диагностических аппаратов на двух разных предприятиях. В Таблице представлены результаты выборочного контроля расхода электроэнергии в квт часах.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| \  \ ***i***  **№** \  **пр.** \ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 13.01 | 14.44 | 14.36 | 13.61 | 13.00 | 12.90 | 13.25 | 12.68 | 13.76 | 14.10 |
| 2 | 15.95 | 15.94 | 14.57 | 14.64 | 14.55 | 15.35 | 14.83 | 15.89 | 14.73 | 15.02 |

Требуется проверить **статистическую гипотезу о равенстве дисперсий** случайного расхода электроэнергии на сборку и наладку 10 медицинских аппаратов на двух разных предприятиях **против альтернативной гипотезы H': *d1*>*d2***. **Математические ожидания** расходов электроэнергии на обоих предприятиях **неизвестны**. Уровень значимости полагается равным 0.15.

**Примечания:**

1.Закон распределения вероятностей случайного расхода электроэнергии предполагается нормальным.

2.Данные статистических выборок независимы.

3.Теоретическое распределение вероятностей расхода электроэнергии на первом предприятии соответствует закону ***N*(13.5; 0.75)**, а на втором - ***N*(15.1; 0.52)**. Этой информацией следует **воспользоваться** **только при проверке правильности** принятого статистического решения **в** **п.3.5** отчетного документа по результатам исследования.

**Дополнительные вопросы**

1. Что называется ошибкой первого рода при принятии статистического решения?
2. По какой формуле вычисляется левая критическая точка левосторонней критической области?
3. Как изменится вероятность ошибки второго рода, если уменьшить вероятность ошибки первого рода?
4. При проверке какой статистической гипотезы статистика критерия подчинена закону распределения вероятностей Стьюдента?

Почему решение о справедливости высказанной гипотезы является статистическим, а не детерминированным?

**Методика оценки результатов выполнения**

контрольной работы по курсу «Математическая статистика»

**Задача № 1**

1. Рассчитаны необходимые выборочные оценки, записано выражение для статистики критерия, рассчитано выборочное значение статистики критерия, рассчитаны критические точки, принято верное статистическое решение, даны ответы на все дополнительные вопросы  **–** 5 баллов.
2. Рассчитаны необходимые выборочные оценки, записано выражение для статистики критерия, рассчитано выборочное значение статистики критерия, рассчитаны критические точки, принято верное статистическое решение, допущена ошибка в расчётах или даны ответы на не все дополнительные вопросы  **– 4** балла.
3. Рассчитаны необходимые выборочные оценки, записано выражение для статистики критерия, рассчитано выборочное значение статистики критерия, рассчитаны критические точки, принято неверное статистическое решение, даны ответы на дополнительные вопросы  **–** 3 балла.
4. Рассчитаны необходимые выборочные оценки, записано выражение для статистики критерия, не рассчитаны выборочное значение статистики критерия и критические точки  **–** 2 балла.
5. Иначе – 1 балл.

|  |
| --- |
| Федеральное государственное автономное образовательное учреждение  высшего образования |
| **«Национальный исследовательский ядерный университет «МИФИ»** |
| ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ |
| КАФЕДРА КИБЕРНЕТИКИ |

**Комплект материалов для оценивания работы на семинарских занятиях по дисциплине**

**«Математическая статистика»**

|  |  |  |
| --- | --- | --- |
| Составитель | Трофимов А.Г. | Доцент, к.т.н. |
| Учебный год | 2018/2019 |  |

1. Поставлена задача проверки соответствия нормативу среднего расхода бензина на 100 км пробега автомобиля в таксомоторном парке. Установленное нормативное значение расхода в Москве, принятое в таксомоторном парке, равно 9 л. В таблице представлены результаты выборочного контроля расхода бензина.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *i* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| *xi* | 8,60 | 9,70 | 7,83 | 8,77 | 9,15 | 9,40 | 9,36 | 8,90 | 10,22 | 7,13 |

Проверяется статистическая гипотеза о равенстве математического ожидания расхода бензина нормативному значению 9 л против альтернативной гипотезы, утверждающей, что в таксомоторном парке имеется значимое превышение нормативного значения расхода бензина. Среднеквадратичное отклонение расхода бензина на 100 км пробега равно 0.75 л. Уровень значимости полагается равным 0.20.

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

2. Рассматривается показатель *X* загрузки складских помещений в каждый месяц календарного года, который зависит от многих случайных факторов и потому является случайной величиной. Закон распределения вероятностей случайной величины *X* неизвестен. Есть предположение, что показатель загрузки имеет нормальное распределение. Для проверки этой гипотезы собраны и сгруппированы многолетние выборочные данные по загрузке в каждый месяц года, которые представлены в таблице (в % от макс. загрузки).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *xi* | 0-0.2 | 0.2-0.4 | 0.4-0.6 | 0.6-0.8 | 0.8-1.0 |
| частота | 17 | 19 | 15 | 14 | 17 |

Используя критерий хи-квадрат, сделать статистический вывод о нормальности распределения показателя X на уровне значимости 0.05.

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

3. В результате статистического исследования доходов предприятий в трёх регионах получены следующие данные:

|  |  |  |  |
| --- | --- | --- | --- |
| Регион | Обследовано предприятий | Средний доход предприятий, млн. руб. | Среднеквадратическое отклонение, млн. руб. |
| Воронежская область | 150 | 100 | 9 |
| Волгоградская область | 120 | 80 | 6 |
| Курская область | 140 | 110 | 7 |

Используя эмпирическое корреляционное отношение, определите тип связи между доходом предприятия и регионом, в котором оно расположено.

4. Исследуется зависимость между ценой спроса и ценой предложения на акции крупнейших предприятий одного из городов РФ. Результаты статистического наблюдения представлены в таблице.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* предприятия п/п | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| цена спроса | 65 | 75 | 85 | 72 | 55 | 60 | 80 | 68 | 85 | 72 |
| цена предложения | 82 | 63 | 93 | 86 | 90 | 76 | 77 | 62 | 93 | 86 |

Рассчитайте ранговый коэффициент корреляции по Спирмену и проверьте его значимость на уровне 0.1. Сравните полученное значение со значением линейного коэффициента корреляции. Какой вывод можно сделать о характере статистической связи между ценой спроса и ценой предложения на акции предприятий?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

5. Проводится анализ связи между результатами психологического тестирования на уровень интеллекта (по 10-балльной системе) и уровнем заработной платы выпускников вузов. В результате статистического наблюдения получены следующие данные.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* п/п | 1 | 2 | 3 | 4 | 5 | 7 | 8 | 9 |
| баллы | 6 | 8 | 6 | 4 | 8 | 9 | 7 | 6 |
| з/п | 12.2 | 18.6 | 15.6 | 16.2 | 25.5 | 17.8 | 14.2 | 12.6 |

Используя корреляционное отношение, определить, есть ли значимое влияние уровня интеллекта, оцениваемого в результате психологического тестирования, и заработной платой выпускников вузов. Сравнить значение корреляционного отношения и коэффициента корреляции. Какой вывод можно сделать о характере связи между рассматриваемым признаками?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

6. Проводится анализ связи между результатами психологического теста на уровень интеллекта и годовыми оценками подростков по алгебре. В результате статистического наблюдения получены следующие данные.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* п/п | 1 | 2 | 3 | 4 | 5 | 7 | 8 | 9 | 10 |
| оценки | 2 | 3 | 3 | 4 | 4 | 5 | 3 | 2 | 4 |
| баллы | 4 | 5 | 7 | 6 | 8 | 8 | 7 | 5 | 7 |

Построить простейшее линейное уравнение регрессии баллов, полученных в результате тестирования, на оценки по алгебре. Является ли построенная линейная модель значимой? Можно ли утверждать, что между рассматриваемыми признаками есть линейная корреляционная связь?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

**Методика оценки результатов работы**

на семинаре по курсу «Основы теории информации и кодирования» за 5 семестр

Оценка знаний студента на семинарском занятии преподавателем определяется следующими основными критериями:

1. Знание учебного материала в соответствии с учебной программой дисциплины – 1 балл.

2. Степень проявления творчества и самостоятельности при решении предлагаемого фрагмента задачи – 1 балл.

3. Доказательность и убедительность правильности своей позиции – 1 балл.

4. Наличие конспекта лекций и его отработка во время самостоятельной работы – 1 балл.

5. Активность на семинарском занятии (умение и стремление задавать вопросы, участие в дискуссии) – 1 балл.

Максимальный балл за работу на семинаре – 5 баллов.

В зависимости от набранных баллов оценки выставляются

в соответствии со следующей таблицей:

|  |  |
| --- | --- |
| 5 | «отлично» |
| 4 | «хорошо» |
| 3 | «удовлетворительно» |
| 0 – 2 | «неудовлетворительно» |

|  |
| --- |
| Федеральное государственное автономное образовательное учреждение  высшего образования |
| **«Национальный исследовательский ядерный университет «МИФИ»** |
| ИНСТИТУТ ИНТЕЛЛЕКТУАЛЬНЫХ КИБЕРНЕТИЧЕСКИХ СИСТЕМ |
| КАФЕДРА КИБЕРНЕТИКИ |

**Вопросы к экзамену по дисциплине**

**«Математическая статистика»**

|  |  |  |
| --- | --- | --- |
| Составитель | Трофимов А.Г. | Доцент, к.т.н. |
| Учебный год | 2018/2019 |  |

Экзаменационный билет включает два теоретических вопроса и две задачи.

**ВОПРОСЫ К ЭКЗАМЕНУ**

1. Понятие статистической гипотезы. Виды гипотез. Примеры (содержательные постановки задач и формулировки основной и альтернативной гипотез). Проверка статистической гипотезы. Общая постановка задачи. Статистика критерия. Принцип построения, требования к свойствам статистики критерия, графическая иллюстрация. Область допустимых значений и критическая область для статистики критерия. Уровень значимости. Типы критических областей, иллюстрация на рисунке. Уравнения для расчета критических точек. Характеристика *p-value*.
2. Проверка статистической гипотезы. Общая постановка задачи. Решающее правило при принятии статистического решения. Статистика критерия. Метод доверительных интервалов для проверки статистической гипотезы. Уровень значимости. Ошибки принятия статистического решения, определения и расчетные выражения. Мощность критерия. Функция мощности критерия. Примеры, графическая иллюстрация.
3. Статистические гипотезы о математическом ожидании случайной величины при известной и неизвестной дисперсии. Содержательный пример. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
4. Статистические гипотезы о дисперсии случайной величины при известном и неизвестном математическом ожидании. Содержательный пример. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистики критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
5. Статистическая гипотеза о равенстве математических ожиданий случайных величин при известных и неизвестных дисперсиях. Содержательный пример. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
6. Статистические гипотезы о равенстве дисперсий случайных величин при известных и неизвестных математических ожиданиях. Содержательный пример. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистики критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
7. Критерий согласия Пирсона (критерий ). Постановка задачи: основная и альтернативная гипотезы, свойства выборок, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
8. Критерий согласия Колмогорова. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
9. Критерий знаков. Постановка задачи: основная и альтернативная гипотезы, свойства выборки, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Предельные свойства закона при . Расчет критических точек. Решающее правило. Графическая иллюстрация.
10. Модификация критерия согласия Пирсона (критерия ) для проверки статистической гипотезы об однородности двух выборок. Постановка задачи: основная и альтернативная гипотезы, свойства выборок, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
11. Модификация критерия Колмогорова для проверки статистической гипотезы об однородности двух выборок. Постановка задачи: основная и альтернативная гипотезы, свойства выборок, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
12. Проверка статистической гипотезы о нормальном законе распределения вероятностей случайной величины на основе анализа коэффициента асимметрии и эксцесса. Постановка задачи: основная и альтернативная гипотезы, свойства выборок, уровень значимости. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Расчет критических точек. Решающее правило. Графическая иллюстрация.
13. Понятие факторного и результативного признака. Виды связей между признаками. Анализ статистической взаимосвязи между двумя качественными признаками. Таблица сопряжённости. Принцип построения. Проверка значимости связи. Постановка задачи: основная и альтернативная гипотезы. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
14. Понятие факторного и результативного признака. Виды связей между признаками. Анализ статистической взаимосвязи между качественным факторным и количественным результативным признаками. Виды дисперсий в совокупности, разделенной на части. Правило сложения дисперсий. Эмпирический коэффициент детерминации. Эмпирическое корреляционное отношение. Расчётные формулы и их интерпретация. Оценка степени тесноты связи. Примеры.
15. Понятие факторного и результативного признака. Виды связей между признаками. Анализ статистической взаимосвязи между качественным факторным и количественным результативным признаками. Однофакторный дисперсионный анализ. Постановка задачи: основная и альтернативная гипотезы, предположения о свойствах выборок. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
16. Понятие факторного и результативного признака. Виды связей между признаками. Анализ корреляционной связи между двумя количественными признаками. Функция регрессии. Вывод уравнения регрессии для случая двумерного нормального распределения. Точечная оценка параметров функции регрессии по выборке из двумерного нормального распределения.
17. Анализ корреляционной связи между двумя количественными признаками. Виды дисперсий в корреляционном анализе. Правило сложения дисперсий. Коэффициент детерминации. Корреляционное отношение. Сравнение коэффициента корреляции и корреляционного отношения как меры оценки корреляционной связи. Расчёт корреляционного отношения для выборки из двумерного нормального распределения.
18. Анализ корреляционной связи между двумя количественными признаками. Проверка значимости корреляционной связи на основе коэффициента корреляции и корреляционного отношения. Постановка задачи: основная и альтернативная гипотезы. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
19. Использование ранговых коэффициентов корреляции для анализа корреляционной связи. Ранговые коэффициенты корреляции по Спирмену и по Кендаллу. Сравнение с линейным коэффициентом корреляции. Проверка значимости ранговой корреляционной связи. Постановка задачи: основная и альтернативная гипотезы. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
20. Задачи регрессионного анализа. Постановка задачи оценивания параметров регрессионной модели. Метод наименьших квадратов. Простейшая линейная регрессионная модель. Расчёт точечных оценок параметров простейшей линейной регрессионной модели. Интервальные оценки параметров простейшей линейной регрессионной модели (без вывода). Характеристики регрессионной модели: точность, значимость, адекватность.
21. Задачи регрессионного анализа. Линейная регрессионная модель общего вида. Постановка задачи оценивания параметров регрессионной модели. Метод наименьших квадратов. Расчёт точечных оценок параметров линейной регрессионной модели общего вида. Система нормальных уравнений Гаусса (без вывода). Интервальные оценки параметров линейной регрессионной модели общего вида (без вывода). Характеристики регрессионной модели: точность, значимость, адекватность. Проверка значимости линейной регрессионной модели общего вида.
22. Анализ множественных статистических взаимосвязей. Множественный коэффициент корреляции. Проверка значимости множественного коэффициента корреляции. Постановка задачи: основная и альтернативная гипотезы. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
23. Анализ множественных статистических взаимосвязей. Коэффициент конкордации. Проверка значимости коэффициента конкордации. Постановка задачи: основная и альтернативная гипотезы. Статистика критерия: расчетная формула, область допустимых значений, критическая область, закон распределения вероятностей. Решающее правило. Графическая иллюстрация.
24. Анализ множественных статистических взаимосвязей. Многомерная линейная регрессионная модель общего вида. Постановка задачи оценивания параметров регрессионной модели. Метод наименьших квадратов. Расчёт точечных оценок параметров многомерной линейной регрессионной модели общего вида. Система нормальных уравнений Гаусса (без вывода). Пример расчёта параметров простейшей линейной регрессионной модели для случая двух факторных переменных.
25. Постановка задачи кластерного анализа. Понятие кластера. Задача отнесения объекта к кластеру. Задача кластерного анализа как задача оптимизации. Мера качества кластеризации. Кластерная дисперсия. Сравнительный анализ методов кластеризации, их преимущества и недостатки.
26. Постановка задачи кластерного анализа. Агломеративные методы кластеризации. Методы агломерации кластеров. Виды расстояний между кластерами. Метод Уорда. Статистическое расстояние.
27. Постановка задачи кластерного анализа. Агломеративные методы кластеризации. Визуальное представление результата агломеративной кластеризации. Дендрограмма: принцип и пример построения. Критерии останова процедуры агломерации кластеров.

**ЗАДАЧИ К ЭКЗАМЕНУ**

1. Исследуется качество производства деталей с использованием нововведенной технологической линии. В таблице представлены выборочные данные по размерам выпущенных деталей в мм.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| I | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *xi* | 2.5 | 2.6 | 2.9 | 2.3 | 2.1 | 2.45 | 2.0 |

Считается, что технологическая линия удовлетворяет требованиям качества, если среднеквадратическое отклонение размера выпускаемых деталей не превышает 0.2 мм.

Можно ли на основе имеющейся выборки данных утверждать, что новая технологическая линия удовлетворяет требованиям качества? Уровень значимости принять 0.1.

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

2. Рассматривается показатель *X* загрузки складских помещений в каждый месяц календарного года, который зависит от многих случайных факторов и потому является случайной величиной. Закон распределения вероятностей случайной величины *X* неизвестен. Есть предположение, что показатель загрузки имеет нормальное распределение. Для проверки этой гипотезы собраны и сгруппированы многолетние выборочные данные по загрузке в каждый месяц года, которые представлены в таблице (в % от макс. загрузки).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *xi* | 0-0.2 | 0.2-0.4 | 0.4-0.6 | 0.6-0.8 | 0.8-1.0 |
| частота | 12 | 16 | 18 | 12 | 11 |

Используя критерий хи-квадрат, сделать статистический вывод о нормальности распределения показателя X на уровне значимости 0.05.

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

3. В результате статистического исследования доходов предприятий в трёх регионах получены следующие данные:

|  |  |  |  |
| --- | --- | --- | --- |
| Регион | Обследовано предприятий | Средний доход предприятий, млн. руб. | Среднеквадратическое отклонение, млн. руб. |
| Воронежская область | 250 | 120 | 12 |
| Волгоградская область | 100 | 130 | 8 |
| Курская область | 150 | 110 | 5 |

Используя эмпирическое корреляционное отношение, определите тип связи между доходом предприятия и регионом, в котором оно расположено.

4. Исследуется зависимость между ценой спроса и ценой предложения на акции крупнейших предприятий одного из городов РФ. Результаты статистического наблюдения представлены в таблице.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* предприятия п/п | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| цена спроса | 83.6 | 83.6 | 30.3 | 13.5 | 13.9 | 26.5 | 18.1 | 28.7 | 19.0 | 19.0 |
| цена предложения | 60.6 | 40.7 | 33.8 | 22.1 | 33.8 | 33.8 | 20.9 | 35.9 | 21.7 | 24.5 |

Рассчитайте ранговый коэффициент корреляции по Спирмену и проверьте его значимость на уровне 0.1. Сравните полученное значение со значением линейного коэффициента корреляции. Какой вывод можно сделать о характере статистической связи между ценой спроса и ценой предложения на акции предприятий?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

5. Проводится анализ связи между результатами психологического тестирования на уровень интеллекта (по 10-балльной системе) и уровнем заработной платы выпускников вузов. В результате статистического наблюдения получены следующие данные.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* п/п | 1 | 2 | 3 | 4 | 5 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| баллы | 6 | 8 | 6 | 4 | 8 | 9 | 7 | 6 | 4 | 9 | 8 | 8 | 7 |
| з/п | 12.2 | 18.6 | 15.6 | 16.2 | 25.5 | 17.8 | 14.2 | 12.6 | 14.2 | 16.5 | 22.8 | 26.2 | 22.6 |

Используя корреляционное отношение, определить, есть ли значимое влияние уровня интеллекта, оцениваемого в результате психологического тестирования, и заработной платой выпускников вузов. Сравнить значение корреляционного отношения и коэффициента корреляции. Какой вывод можно сделать о характере связи между рассматриваемым признаками?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

6. Проводится анализ связи между результатами психологического теста на уровень интеллекта и годовыми оценками подростков по алгебре. В результате статистического наблюдения получены следующие данные.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* п/п | 1 | 2 | 3 | 4 | 5 | 7 | 8 | 9 | 10 |
| оценки | 3 | 4 | 3 | 5 | 4 | 4 | 3 | 2 | 4 |
| баллы | 6 | 7 | 7 | 9 | 8 | 6 | 8 | 5 | 9 |

Построить простейшее линейное уравнение регрессии баллов, полученных в результате тестирования, на оценки по алгебре. Является ли построенная линейная модель значимой? Можно ли утверждать, что между рассматриваемыми признаками есть линейная корреляционная связь?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

7. Имеются следующие данные о суммах активов, кредитных вложениях и вкладах граждан в коммерческие банки одного из регионов РФ.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *№* п/п банка | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| кредитные вложения | 3.1 | 6.5 | 11.4 | 19.6 | 13.2 | 24.9 | 28.2 | 9.2 |
| сумма активов | 5.1 | 11.2 | 18.6 | 30.7 | 20.0 | 31.7 | 33.4 | 16.4 |
| вклады граждан | 0.42 | 0.50 | 0.55 | 0.52 | 0.58 | 0.84 | 0.65 | 0.58 |

Постройте простейшее линейное уравнение регрессии суммы активов коммерческого банка на две переменные: кредитные вложения и вклады граждан. Является ли построенная линейная модель значимой? Можно ли утверждать, что между рассматриваемыми признаками есть линейная корреляционная связь?

Решение задачи должно содержать расчетную формулу для статистики критерия, закон распределения вероятностей статистики, графическую иллюстрацию к построению критической области, значение(я) критической(их) точки(ек) и статистический вывод.

**Методика оценки результатов сдачи экзамена**

по курсу «Математическая статистика» за 5 семестр

Критерии оценки знаний устанавливаются в соответствии с требованиями к профессиональной подготовке, исходя из действующих учебных планов и программ, с учётом характера будущей практической деятельности выпускника.

**«ОТЛИЧНО»** (45-50 баллов) - студент владеет знаниями предмета в соответствии с рабочей программой, достаточно глубоко осмысливает дисциплину; самостоятельно, в логической последовательности и исчерпывающе отвечает на вопрос билета, четко формулирует ответ и решает задачу билета в полном объеме.

**«ХОРОШО»** (35-44 баллов) - студент владеет знаниями дисциплины почти в полном объеме программы (имеются пробелы знаний только в некоторых, особенно сложных разделах); самостоятельно и отчасти при наводящих вопросах дает полноценный ответ на вопрос билета; не допускает серьезных ошибок при решении задачи билета.

**«УДОВЛЕТВОРИТЕЛЬНО»** (30-34 баллов) - студент владеет основным объемом знаний по дисциплине; проявляет затруднения в самостоятельных ответах, оперирует неточными формулировками; в процессе ответов допускаются ошибки по существу вопросов; способен решать задачу билета не в полном объеме.

**«НЕУДОВЛЕТВОРИТЕЛЬНО»** (ниже 30 баллов) - студент не освоил обязательного минимума знаний предмета; не способен ответить на вопрос билета даже при дополнительных наводящих вопросах экзаменатора; не может решить задачу билета.

**Итоговая оценка по курсу выставляется в соответствии**

**со следующей таблицей:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Сумма баллов по дисциплине** | **Оценка по 4-х бальной шкале** | **Зачет** | **Оценка (ECTS)** | **Градация** |
| 90 - 100 | 5 (отлично) | Зачтено | А | Отлично |
| 85 - 89 | 4 (хорошо) | В | Очень хорошо |
| 75 - 84 | С | Хорошо |
| 70 - 74 | 3 (удовлетворительно) | D | Удовлетворительно |
| 65 - 69 |
| 60 - 64 | E | Посредственно |
| Ниже 60 | 2 (неудовлетворительно) | Не зачтено | F | Неудовлетворительно |

**В результате освоения дисциплины студент должен:**

**Знать:**

**З-1** основные понятия и методы математической статистики;

**З-2** методы статистического описания результатов наблюдений и статистического анализа данных.

**Уметь:**

**У-1** формулировать задачи обработки данных в терминах математической статистики;

**У-2** применять методы статистического анализа на реальных данных.

**Владеть:**

**В-1** навыками обработки данных с использованием статистических пакетов.

|  |  |  |  |
| --- | --- | --- | --- |
| **Компетенции** | **Знания (знать)** | **Умения (уметь)** | **Навыки (владеть)** |
| ПК-12: | З-1 | У-1 | В-1, В-2 |
| ПК-16: | З-1, З-2 | У-2 | В-1, В-2 |