Дисциплина: Биостатистика

Курс: 4

Направление: 06.03.01 – Биология

Преподаватель: к.б.н., доцент Каюкова С.Н.

**ЛПЗ на период 16.11.2020 г.**

**Задание 1.**

**Математическое ожидание** характеризует среднее значение случайной величины и определяется по формулам:



В качестве математического ожидания используется «среднее взвешенное значение», причем каждое из значений случайной величины учитывается с «весом», пропорциональным вероятности этого значения.

Физический смысл математического ожидания – среднее значение случайной величины, т.е. то значение, которое может быть использовано вместо случайной величины в приблизительных расчетах или оценках.

**Задание 2.**

**Дисперсия** случайной величины характеризует степень рассеивания (разброса) значений случайной величины относительно ее математического ожидания и определяется по формулам:



**Задание 3.**

Дисперсия случайной величины имеет размерность квадрата случайной величины, поэтому для анализа диапазона значений величины Х дисперсия не совсем удобна. Этого недостатка лишено среднеквадратическое отклонение (СКО), размерность которого совпадает с размерностью случайной величины.

**Среднеквадратическое отклонение** случайной величины X характеризует ширину диапазона значений X и равно:



СКО измеряется в тех же физических единицах, что и случайная величина.

Рассчитайте все показатели, если:

p = 1,2,3,4,5,6,7,8,9,10. Хi – доля p.

**Задание 4.** Наблюдения интегральным методом на двух участках, показали, что на первом из них процент учетных единиц, перешедших межу, оказался равным 54 + - 8%, а на втором – 37 + - 3%. Можем ли мы утверждать на основании полученных результатов, что сезонное развитие изучаемого объекта на первом участке идет с опережением по сравнению со вторым? Приведите соответствующие расчеты.

**Задание 5.** Размер крыльев 25 птенцов по интегральному методу составил (в см):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 6 | 11 | 7 | 8 | 9 |
| 10 | 5 | 9 | 7 | 4 |
| 12 | 7 | 9 | 8 | 6 |
| 5 | 6 | 7 | 13 | 9 |
| 12 | 10 | 9 | 14 | 7 |

Вычислите среднее квадратичное отклонение, среднюю ошибку средней арифметической и коэффициент вариации.

Конспект:

**Тема лекции: 4. Методы биостатистики**

В основе обработки и анализа данных лежат известные математические методы.

Благодаря использованию информационных технологий, в настоящее время этап обработки данных стал наименее трудоемким. На первое место в отношении трудоемкости вышли такие этапы, как освоение статистических пакетов, этап подготовки данных к анализу, этап предварительного анализа данных и этап интерпретации результатов. Все в целом привело к изменениям в технологии обработки и анализа данных.

При этом для выполнения методов обработки медико-биологических данных от пользователя требуется лишь применение статистических методов обработки данных и использования соответствующих пакетов прикладных программ. Доктор, как правило, не требуется углубляться в сложные математические теории, а надо понимать, для чего и каким образом они используются.

На практике для врача обработка и анализ данных сводятся к решению следующих задач: получение представления о основные статистические методы; усвоения пакета прикладных программ; анализа и интерпретация результатов исследований.

Сам анализ данных с использованием статистического пакета (работа с пакетом, сама технология анализа данных) включает в себя следующие этапы:

* планирование исследования
* подготовка данных к анализу
* предварительный анализ данных
* выбор метода анализа и его реализация
* интерпретация результатов
* представление результатов.

**4.1 Планирование исследования**

На начальных этапах исследования нет четкости относительно методов обработки результатов. Поэтому надо предусмотреть возможность использования различных способов обработки и приблизительного сравнения полученных результатов с целью определения, как обрабатывать имеющиеся данные. Приведенная ниже таблица поможет лучше сориентироваться в основных математических методах обработки и анализа данных.

Таблица 1. Математические методы обработки и анализа данных

|  |  |
| --- | --- |
| Источник информации, задача исследования | Методы обработки и анализа |
| Анкеты, вопросы, тесты.  Исследование состояния здоровья, клинические обследования, медицинские записи, регистрации, истории болезни.  Выявление взаимосвязей.  Описание случаев заболеваний, сопоставление с данными прошлых лет, изучение осложнений.  Исследование эффективности различных процедур, изучение связей между процедурами и их последствиями. | Классификация данных.  Создание таблиц и отчетов.  Корреляционный анализ. Факторный анализ. Регрессионный анализ. Дисперсионный анализ.  Методы проверки гипотез. |
| Скрининговые исследования | Дискриминантный анализ, кластерный анализ, методы распознавания образов. |
| Медико-статистические данные.  Исследование заболеваемости, динамика заболеваемости, выявление периодичности. | Методы анализа случайных процессов,  спектральный анализ, математическое моделирование. Методы распознавания образов. Теория массового обслуживания. Параметрический и непараметрический статистический анализ. |
| Лабораторные эксперименты и опыты на животных. | Статистические методы планирования экспериментов. Регрессионный анализ, дисперсионный анализ, многомерный статистический анализ, методы математического моделирования. |
| Клинические исследования.  Сравнительные лечебные исследования, анализ выживаемости и наследственности с учетом принадлежности пациента к определенной группе, изучение дозировки препаратов.  Разработка методов диагностики. | Дисперсионный анализ, регрессионный анализ, дискриминантный анализ, методы проверки гипотез. Математическое моделирование процессов. Дискриминантный анализ, кластерный анализ, методы распознавания образов. |
| Клинические лабораторные данные.  Сбор, хранение и передача клинической информации, анализ качества и надежности лабораторных исследований, уход за пациентом. | Статистический анализ. Дисперсионный анализ. Регрессионный анализ. Последовательный анализ Вальда. |

**4.2 Подготовка данных к анализу**

Целью этого этапа является приведение данных к виду, позволит провести следующую их обработку, и предварительное формирование представления о типе (структуру) анализируемых данных.

Обычно во время проведения медицинского исследования стараются учесть максимальное количество характеристик, которые существенны при анализе исследуемому вопросу. Исследования, как правило, состоит из нескольких серий наблюдений, во время которых в одинаковых условиях регистрируются параметры отдельных объектов (например, больных определенное заболевание). Имея дело с серией наблюдений, надо пытаться представить их в простой форме, которая позволила бы непосредственно или путем следующих вычислений сделать из них выводы.

Все данные целесообразно свести к единой таблицы, в которой по строкам расположены различные объекты наблюдения (например, больные), а по столбцам параметры (например, температура, частота сердечных сокращений, артериальное давление и др.). В этой таблице объекты могут быть объединены в несколько групп в соответствии с общим признакам (по возрасту, полу и т.д.).

Заметим, что введенные однажды данные могут быть обработаны с помощью различных методов.

Важным этапом в подготовке данных к анализу является визуализация, или просмотр данных. Чтобы выяснить, какие методы анализа нужно использовать к данным и насколько Вас удовлетворяют полученные результаты после выполнения статистических процедур, надо иметь возможность наглядно представить себе эти данные и результаты. Ведь изучение таблиц менее удобная процедура, чем представление данных в виде графиков и диаграмм. Графические образы в медицинских научных разработках помогают наблюдать за тенденциями изменений, выявлять сложные взаимодействующие факторы и упрощает сопоставление данных.

Таким образом, использование графиков облегчает предварительный анализ информации, поэтому целесообразно строить различные графики для лучшего понимания полученного экспериментального материала позволяет получить общее представление об особенностях и закономерностях обрабатываемых. Результаты использования статистических процедур, как правило, подаются в наглядном графическом виде.

**4.3 Предварительный анализ данных**

Целью этого этапа является формирование представления о типе (структуру) анализируемых данных, и предварительный выбор методов анализа. Этот этап включает: уточнение структуры данных и разбиения их на группы; расчет основных статистических характеристик, выявление различий между группами данных, определение взаимосвязей между параметрами, определение эмпирических законов распределения, которым подчинены данные.

**4.4 Выбор и реализация метода анализа**

В современных прикладных пакетах данные достаточно просто обрабатываются с помощью различных процедур, с тем чтобы потом можно было выбрать метод, который дает наилучший результат.

Использование определенного статистического метода определяется общей целью исследования. Например, если необходимо оценить степень влияния известных факторов на величину, измеряемую, используют дисперсионный и / или регрессионный анализы. Если из множества факторов воздействия необходимо выявить несколько ведущих - используют регрессионный и факторный анализы. Для оценки особенностей явления меняется со временем, используют корреляционный и спектральный анализы. Чтобы разбить совокупность объектов, изучаемых на группы "похожих" случаев, используют кластерный анализ, дискриминатний анализ.