

Введение в предмет «Биологическая статистика».

Доцент модуля медицинской
биофизики и биостатистики
Аймаханова Айзат Шалхаровна

План лекции:

1. Предмет и задачи биологической статистики.
2. Данные:
 - а) Виды данных
 - б) Шкалы измерения данных
 - в) Графическое представление данных
3. Выборочный метод.

XXI век – век доказательной медицины.

Большую роль в медицине играют методы точных наук, в первую очередь статистики.

В чем суть доказательной медицины?

Д. м. требует, чтобы при решении задач медицинской практики использовались результаты лучших клинических испытаний, в надежности и доброкачественности которых нет никаких сомнений.

Д.м. значительно сокращает и облегчает перенос научных достижений в медицинскую практику.

Практикующий врач должен понимать современный статистический язык и знать методологию современной медицины. В ином случае врач не сможет эффективно пополнять свои знания при чтении современных медицинских публикаций.

Внедрение в практику принципов д.м. диктует современному выпускнику мед. вузов необходимость понимания биостатистики.

Статистика

- Это наука, позволяющая увидеть закономерности в хаосе случайных данных, выделить устойчивые связи в них, определить действия с тем, чтобы увеличить долю правильно принятых решений среди всех принимаемых.

Статистика

- Это сбор, группировка, систематизация, представление, анализ и интерпретация данных (результатов наблюдений).

Это изучение генеральных совокупностей и их изменчивости.

Статистика

- промышленная,
- сельскохозяйственная,
- коммунальная,
- судебная,
- и другие,
- биологическая – изучающая вопросы связанные с медициной и биологией.

Биологическая статистика

- система знаний о правилах планирования и анализа результатов натурных и экспериментальных исследований биологических объектов.

Биостатистика-как дисциплина

Предмет – основные положения и методы биологической статистики.

Цель дисциплины – научить правильно количественно оценивать процессы, происходящие в живой природе.

Задачи дисциплины – познакомить и научить пользоваться основными статистическими критериями для оценки биологических совокупностей.

Структура данных



Качественные данные

Номинальные или категориальные
(классификационные):

Данные разбиты на категории, названия которых отражают значения признака или классификационный номер.

Данные этого типа нельзя измерять или упорядочивать по значению, над ними нельзя выполнять арифметические операции.

Примеры: пол, семейное положение, этническая группа, статус занятости, группа крови.

Качественные данные

Порядковые (ранговые):

Данные относятся к непересекающимся категориям, в зависимости от степени проявления у них анализируемого признака. Категории (градации, уровни) логически упорядочены в соответствии со шкалой относительной значимости значений.

Над данными нельзя выполнять арифметические операции.

Примеры: стадии заболевания, выраженность боли (шкала оценки боли), классы общества.

Качественные данные

Бинарные или дихотомические:

Номинальные или категориальные данные, которые распределяются по двум непересекающимся категориям.

Примеры: мертвый / живой, излеченный / не излеченный.

Количественные (числовые) данные

Дискретные числовые :

Данные могут иметь только строго фиксированные числовые значения, обычно целые. Значения данных изменяются скачкообразно, и никаких промежуточных значений не существует.

Для этих данных важны и величина, и порядок.

Примеры: количество дней, число пациентов.

Количественные (числовые) данные

Непрерывные числовые :

Данные могут принимать любое числовое значение из заданного диапазона (отрезка). Количество значений бесконечно, значения могут содержать дробную составляющую, зависящую от точности измерений.

Примеры: температура, вес, рост, давление.

Графическое представление данных

Применяется для:

- Распознавания структуры и характерных особенностей данных;
- Обнаружения ошибок в данных;
- Исследования взаимосвязи между случайными величинами;
- Оценки соответствия данных выбранным моделям;
- Обнаружения новых явлений;
- Выявления необходимости в корректировке данных, а также определения способа ее осуществления (преобразование данных, сбор дополнительных данных или изменение условий эксперимента).

Графическое представление дискретных данных

- Таблица
- Столбиковая диаграмма
- Круговая диаграмма
- Точечный график
- и т.д.

Графическое представление непрерывных данных

- Группированная выборка
- Гистограмма
- Диаграмма «Стебель с листьями»
- и т.д.

Множество объектов, характеризуемых некоторым качественным или количественным признаком, называется **статистической совокупностью**.

Статистическая совокупность, состоящая из всех объектов, которые (по крайней мере, теоретически) подлежат обследованию, называется **генеральной статистической совокупностью**.

Статистическая совокупность, состоящая из некоторого количества объектов, случайным образом отобранных из соответствующей генеральной совокупности, называется **выборочной совокупностью** или просто **выборкой**.

Метод статистического исследования, состоящий в том, что на основе изучения выборочной совокупности делается заключение о всей генеральной совокупности, называется **выборочным методом**.

Основные требования к выборке – хорошо представлять генеральную совокупность, то есть **быть представительной** (репрезентативной).

Репрезентативность:

- ✓ случайностью выбора объектов из генеральной совокупности, когда каждому из них обеспечивается одинаковая возможность быть отобранным;
- ✓ независимостью результатов наблюдений в выборе;
- ✓ правильным определением объема выборки с учетом всех конкретных условий.

характеристики выборки:

✓ выборочная средняя:

$$\bar{x}_v = \frac{1}{n} \sum_{i=1}^k n_i x_i$$

✓ выборочная дисперсия:

$$D_B = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x}_v)^2$$

✓ выборочное среднеквадратичное отклонение:

$$\sigma_B = \sqrt{D_B}$$

Точечные и интервальные оценки параметров совокупности.

Точечной называют оценку, которая определяется одним числом.

Интервальная - двумя числами.

✓ доверительная вероятность:

$$P=0,95 \text{ (95\%),}$$

✓ уровень значимости:

$$\alpha = 1 - 0,95 = 0,05 \text{ (5\%)}$$

параметров ген. совокупности:

✓ генеральная средняя:

$$\overline{x}_G \approx \overline{x}_v$$

Интервальные оценки параметров совокупности.

Интервал, который с заданной вероятностью (надежностью) покрывает неизвестный параметр, называется **доверительным интервалом**.

Вероятности, признанные достаточными для уверенного суждения о генеральных параметрах на основании известных выборочных показателей, называются **доверительными вероятностями**.

Доверительный интервал для выборочного среднего значения:

$$\left(\bar{x} - t_{\alpha} S_X; \bar{x} + t_{\alpha} S_X \right)$$

где t_{α} - коэффициент Стьюдента
стандартная ошибка среднего:

$$S_X = \frac{\sigma}{\sqrt{n}}$$

Литература:

1. Лукьянова Е.А. Медицинская статистика. – М: Изд. РУДН, 2002.
2. Медик В.А., Токмачев М.С., Фишман Б.Б. Теоретическая статистика//Статистика в медицине и биологии. В 2-х томах / Под. ред. проф. Ю.М Комарова. Т.1. – М.: Медицина, 2000.
4. Гланц С. Медико-биологическая статистика. – М.: Практика, 1999.
5. Рокицкий П.Ф. Биологическая статистика. – Высшая школа, 1973.

**Спасибо за
внимание.**