Математическая статистика и её роль в медицине и здравоохранении.

Математическая статистика – наука о математических методах систематизации и использования статистических данных для научных и практических выводов.

В результате применения статистического метода мы получаем оценку вероятности того или иного предположения. Кроме того каждый статистический метод основан на собственной математической модели и результат его правильный настолько, насколько эта модель соответствует действительности.

Случайная величина- величина, которая в результате испытания может примет одно и только одно возможное значение наперед не известное и не зависящее от случайных величин, которые заранее не могут быть учтены.

Обозначение случайной величины: X,Y,Z.

Значения случайной величины : $x_1, x_2, x_{3,\dots,} y_1, y_2, y_{3\dots}$

Дискретной случайной величиной называют такую величину, множество значений которой либо конечное, либо бесконечное, но *счетное.*

Счетное множество- это бесконечное множество, элементы которого возможно пронумеровать натуральными числами.

Примеры дискретной случайной величины:

- количество пациентов с диагнозом « грипп »,
- число патронажей на дому в день,
- количество поставщиков лекарственных препаратов в аптеку,
- пульс; рост, вес, артериальное давление…

Непрерывной случайной величиной называют случайную величину, которая может принимать все значения из некоторого конечного или бесконечного промежутка. Множество возможных значений непрерывной случайной величины бесконечно и несчетно. Возникает при измерениях.

Примеры непрерывных случайных величин:

- расстояние между населенными пунктами;
- показатели крови (холестерин, гемоглобин, сахар…).

Закон распределения дискретной случайной величины

x_i	x_1	x_2	x_3	$\cdots$	x_k
p_i	p_1	p_2	p_3	$\cdots$	p_k

Таблица задает закон распределения случайной величины X, если выполняется равенство:

$$\sum_{i=1}^{k} p_i = p_1 + p_2 + \cdots + p_k = 1$$

Генеральная статистическая совокупность- совокупность всех исследуемых объектов (бесконечное большая величина).

Выборочная совокупность или выборка - множества объектов, случайно отобранных из генеральной совокупности.

Число наблюдений в совокупности называется ее **объемом.**

N- объем генеральной совокупности.

n- объем выборки.

Варианта x_i - значения случайной величины.

Частота встречаемости n_i – означает, сколько раз встретилось значение x_i.

Вариационный ряд - выборка, представляющая собой неубывающую числовую последовательность.

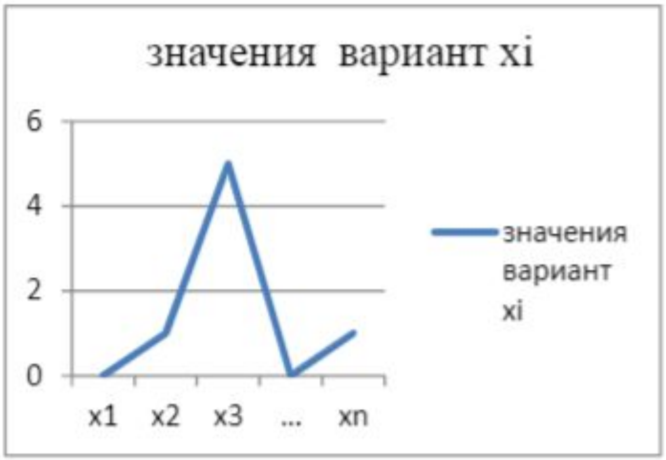
Статистическое распределение (статистический ряд) записывают в виде таблицы:

x_i	x_1	x_2	x_3	…	x_k
n_i	n_1	n_2	n_3	…	n_k

x_i- варианты,

n_i- частота встречаемости варианты x_i.

Для графического изображения статистического дискретного ряда на координатной плоскости откладываются точки (x_i, n_i) и соединяются отрезками, образуя ломаную - **полигон частот**.



Выборочное распределение - записывают в виде таблицы.

x_i	x_1	x_2	x_3	$\dots$	x_k
p_i	$p_1 = \frac{n_1}{n}$	$p_2 = \frac{n_2}{n}$	$p_3 = \frac{n_2}{n}$	$\dots$	$p_k = \frac{n_k}{n}$

p_i- относительные частоты встречаемости значения x_i

$n = n_1 + n_2 + \cdots + n_k$- объем выборки

Основные числовые характеристики случайной величины

Размах выборки - разность между максимальным и минимальным значением вариант.

Медиана (Ме)- это серединная, центральная варианта, делящая вариационный ряд пополам на две части.

Например, если число наблюдений составляет 33, медианой будет варианта, занимающая 17-е ранговое место, так как в обе стороны от нее находится по 16 наблюдений . В ряде с четным числом наблюдений за медиану принимается полусумма в центре находящихся двух величин.

Мода (Мо)- это чаще всего встречающаяся или наиболее часто повторяющаяся величина признака. При приближенном нахождении моды в простом (на сгруппированном) ряде, она определяется как варианта с наибольшим количеством частот.

Математическое ожидание (выборочное среднее)- среднее арифметическое выборки.

Если задано выборочное распределение:

$$\overline{\mathrm{X}} = \mathrm{M(X)} = \sum_{i=1}^{k} x_i \cdot p_i = x_1 \cdot p_1 + x_2 \cdot p_2 + \cdots + x_k \cdot p_k$$

Если задано статистическое распределение:

$$\overline{\mathrm{X}} = \mathrm{M(X)} = \frac{\sum_{i=1}^{k} x_i \cdot p_i}{n} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \cdots + x_k \cdot p_k}{n}$$

Практическое применение средних величин

1.Для оценки состояния здоровья, например, параметров физического развития (средний рост, средний вес, средний объем жизненной емкости легких и т.д.) соматических показателей (средний уровень сахара в крови, средний пульс и т.д.)

2. Для оценки организации работы лечебно-профилактических и санитарно-противоэпидемических учреждений, а также деятельности отдельных врачей и других медицинских работников (средняя длительность пребывания больного на койке, среднее число посещений на 1 ч приема).

3. Для оценки состояния окружающей среды.

Дисперсия («рассеяние») случайной величины- мера разброса случайной величины, равная математическому ожиданию квадрата отклонения случайной величины от ее математического ожидания. Чем больше разброс, тем больше дисперсия.

Если случайная величина задана статистическим рядом:

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

Если величина задана выборочным распределением:

D(X)=$\sum_{i=1}^{k}(x_i-\bar{x})^2\cdot p_i$

Также можно воспользоваться формулой:

D(X)=M(X^2)-$[$M(X)$]^2$, где M(X^2)=$\sum x_i^2 p_i$

Среднее квадратическое отклонение:

$\sigma(X)=\sqrt{D(X)}$

При помощи квадратического отклонения можно установить степень типичности средней, пределы рассеяния ряда, пределы колебаний вокруг средней отдельных вариант.

Применение среднего квадратического отклонения дает возможность оценки и сравнения разнообразия нескольких однородных рядов распределения, так как σ- величина именная, выражается абсолютным числом в единицах изучаемой совокупности (см, кг,мл/л, и т.д)

Примеры решения задач:

1.*Статистическое распределение случайной величины представлено в таблице. Вычислите объем выборки и размах, моду (Мо) и медиану (Ме).*

X_i	1	2	5	6	8	10	12	13	15
P_i	2	3	3	5	6	4	4	2	1

Решение:

1.Объем выборки - сумма n_i

n=2+3+3+5+6+4+4+2+1=30

2. Размах выборки: Δ=15-1=14.

3. Модой является варианта x=8, Мо=8.

4. Медианой является полусумма 15 и 16 вариант $x_{15} = x_{16} = 8$, значит Ме=8

Ответ: n=30; Δ= 14; Мо=8; Ме=8.