

**Законы распределения
случайных величин.**

**Методы математической
статистики: проверка
статистических гипотез.**

Лекция 6

План

- 1. Случайные величины
- 2. Закон распределения случайных величин
- 3. Статистическая гипотеза
- 4. Статистические критерии проверки гипотез

1. Случайные величины.

Случайная величина – это такая величина, которая в результате опыта может принять то или иное значение, причем неизвестно заранее какое именно.

Примеры: число попаданий при трех выстрелах; число вызовов, поступавших на телефонную станцию за сутки

Различают:

Одномерные Случайные величины

- 1) дискретными
- 2) непрерывными

Многомерные Случайные величины

- 1) дискретными
- 2) непрерывными
- 3) и их комбинациями

Непрерывная случайная величина

Случайная величина называется ***непрерывной***, если все её возможные значения полностью заполняют какой-либо конечный или бесконечный интервал числовой оси.

Например: масса новорожденного, температура, рН крови, концентрация молекул...

Дискретная случайная величина

Величина, принимающая
отдельные, изолированные
возможные значения с
определенными вероятностями
называется *дискретной
случайной величиной*.

*Например: число студентов в аудитории, число людей
в очереди, транспорте, число книг в библиотеке...*

Случайную величину принято обозначать : X,Y,Z...

$$X = \{1,2,3,4,5,6\}$$

Например, X - количество очков, выпавших на игральной кости:

$$X = \{x_1, x_2, x_3, x_4 \dots x_n\}$$

2. ЗАКОН РАСПРЕДЕЛЕНИЯ

дискретной случайной величины

ЗАКОНОМ РАСПРЕДЕЛЕНИЯ

дискретной случайной величины

называют всякое соответствие

между всеми возможными

значениями случайной величины и

вероятностями, с которыми она

принимает то или иное значение

Закон распределения ДСВ можно представить в виде таблицы (матрицы)

X	X₁	X₂	X₃	...	X_n
P	P₁	P₂	P₃	...	P_n

Причем,
$$\sum_{i=1}^n p_i = 1$$

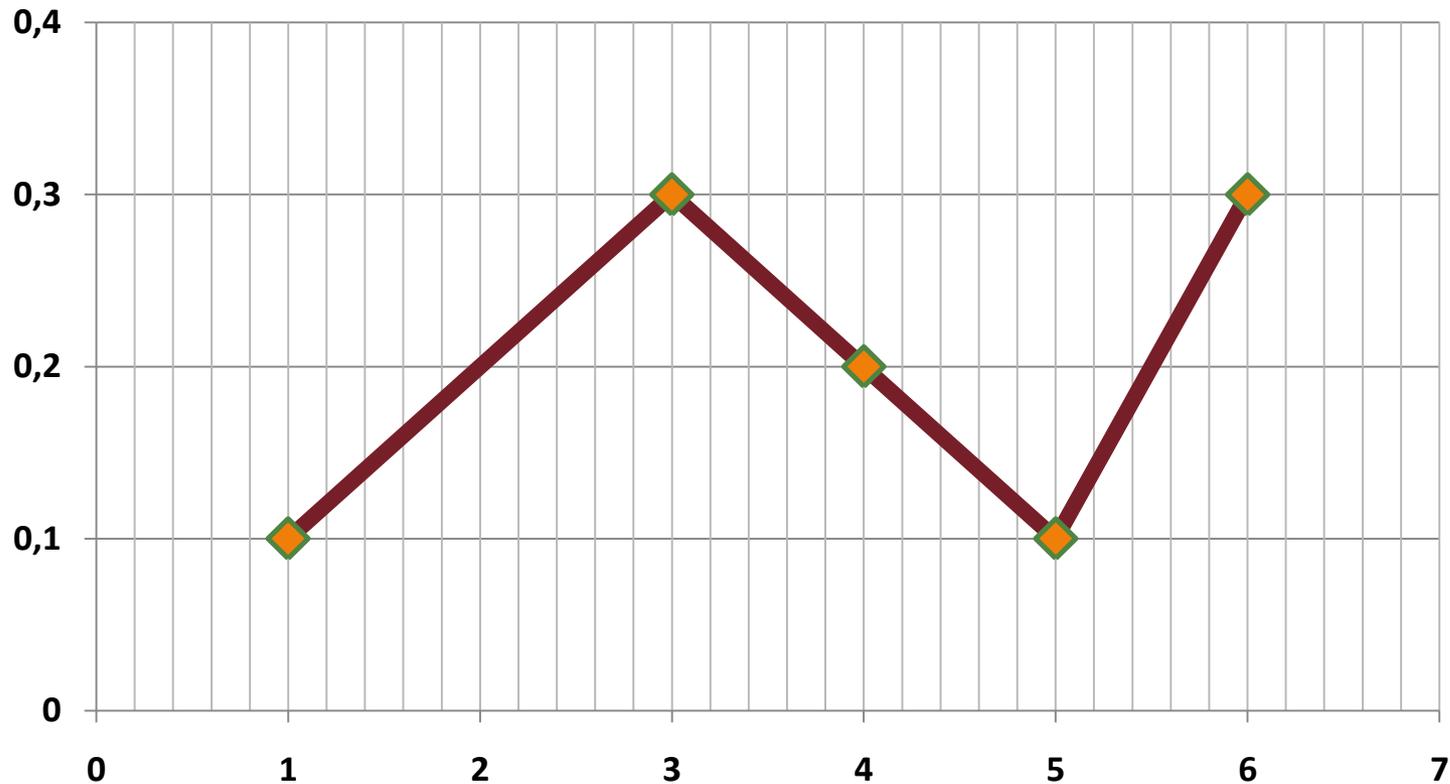
Закон распределения ДСВ- графический

- Множество точек $M(x_i, p_i)$, соединённых ломаной линией, называется **многоугольником распределения**
- Например, закон распределения ДСВ задан в виде таблицы (матрицы)

x	1	3	4	5	6	Σ
p	0,1	0,3	0,2	0,1	0,3	1

ПОСТРОИТЬ МНОГОУГОЛЬНИК РАСПРЕДЕЛЕНИЯ

Многоугольник распределения
случайной величины X



СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ

Гипотеза -это научное предположение, допущение, истинное значение которого неопределенно.

Статистической гипотезой называют гипотезу о виде неизвестного распределения или о параметрах известных распределений.

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ В МЕДИЦИНЕ

- Сравнительная проверка и оценка эффективности лечения, продолжительности болезни и восстановительного периода, тяжести заболевания;
 - Сравнение лечебных методик;
- Сравнение различных характеристик препаратов и медицинской техники;
- сравнение различных мер профилактики.

Пример: сравнения двух методик лечения некоторого заболевания

1 группа (методика №1)



2 группа (методика №2)



Для каждого пациента фиксировалось количество процедур, после которого достигался положительный эффект.

Число процедур	3	4	5	6	7	8	9	10	11	12	\bar{X}_{cp}	σ
1-я группа	1	2	3	7	4	2	1	-	-	-	6,05	1,45
2-я группа	1	-	2	3	6	3	2	1	1	1	7,75	2,01

выборочное среднее в первой группе меньше, чем во второй

Можно ли утверждать, что и в генеральных совокупностях, из которых извлечены 2 выборки испытуемых, наблюдается такое же различие между генеральными средними значениями?

СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ:

- 1. Нулевая гипотеза (H_0)** – различия между выборочными характеристиками случайны, недостоверны.
- 2. Альтернативная гипотеза (H_1)** – различия между выборочными характеристиками достоверны, т. е. реально наблюдаются между генеральными параметрами в генеральных совокупностях, из которых извлечены выборки.

Проверка гипотез

Выдвинутая гипотеза может оказаться правильной или неправильной. Поэтому она проверяется.

ПРИ СТАТИСТИЧЕСКОЙ ПРОВЕРКЕ МОГУТ БЫТЬ ДОПУЩЕНЫ ОШИБКИ!

Гипотеза H_0	Принимается	Отвергается
Верна	Правильное решение	Ошибка I рода
Не верна	Ошибка II рода	Правильное решение

Ошибка первого рода - отвергается правильная гипотеза

Вероятность совершить ошибку первого рода называют **уровнем значимости (α)**

В биологии и медицине уровень значимости принимают не выше 0,05. Это означает, что в 5 случаях из 100 (5%) мы рискуем допустить ошибку первого рода.

Ошибка второго рода - принимается неправильная гипотеза

Значимость ошибки второго рода
обозначают символом β .

*Тогда вероятность не допустить ошибку
II рода будет равна $(1 - \beta)$; ее называют
мощностью критерия (или функцией
мощности)*

СТАТИСТИЧЕСКИЙ КРИТЕРИЙ

- строгое математическое правила, по которому принимается или отвергается та или иная статистическая гипотеза с известным уровнем значимости.

Критерии согласия

- Критерии, определяющие степень соответствия известному распределению теоретическому распределению

Критерии различия

- Критерии, определяющие достоверность различий 2-х выборок по определенным характеристикам (среднему или дисперсии)

Параметрические критерии

- *Критерии, которые рассчитываются на основе параметров выборочной совокупности (Критерий Стьюдента, F-критерий Фишера)*

Непараметрические критерии

- *Критерии, которые рассчитываются на основе частоты встречаемости или на основе рангов (Критерий согласия Пирсона, Критерий «Хи-квадрат»)*

ОБЩАЯ СХЕМА ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

- 1) Формулируются гипотезы H_0 и H_1 .
- 2) Выбирается уровень значимости критерия α .
- 3) Вычисляется $K_{\text{набл}}$ (статистика критерия).
- 4) Находится $K_{\text{крит}}$ критическое (граничное) значение критерия.
- 5) Значение $K_{\text{набл}}$ критерия сравнивается с $K_{\text{крит}}$ и по результатам сравнения делается вывод: принять гипотезу или отвергнуть.

$$K_{\text{набл}} < K_{\text{крит}}$$

гипотеза H_0 принимается на заданном уровне значимости α

$$K_{\text{набл}} > K_{\text{крит}}$$

гипотеза H_0 отклоняется, принимается гипотеза H_1 при данном уровне значимости α .