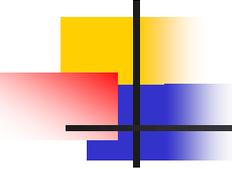


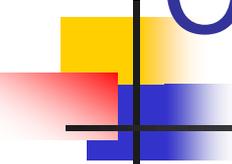
# Выборочная дисперсия



# План:

---

1. Основные задачи статистики
2. Выборочный метод
3. Статистические оценки параметров распределения
4. Выборочная дисперсия



# Основные задачи статистики

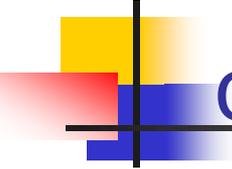
---

## **Задачи статистики:**

- Указать способы сбора и группировки статистических данных, полученных в результате наблюдений или экспериментов
- Анализ статистических данных в зависимости от целей исследования

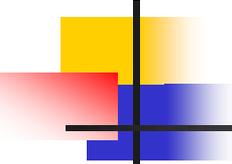
## **Цели исследования:**

- Оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения; оценка зависимости случайной величины от одной или нескольких случайных факторов и др.
- Проверка статистических гипотез о виде неизвестного распределения или о величине параметров распределения



## Статистическая обработка данных осуществляется по следующим этапам:

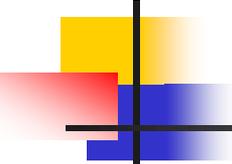
1. Планирование процесса сбора данных
2. Подготовка данных для анализа
3. Анализ данных
4. Формулировка выводов на основе анализа данных



# Терминология

---

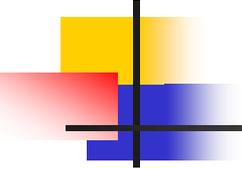
- *Генеральная совокупность (population)* это совокупность всех объектов
- *Выборочная совокупность* или просто *выборка (sample)* это совокупность случайно отобранных из генеральной совокупности объектов
- *Объемом* совокупности называют число объектов этой совокупности (*объем генеральной совокупности - population size, объем выборки - sample size*)



# Выборочный метод

---

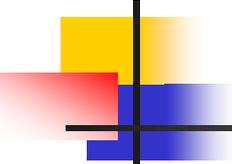
- **Повторной** называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность
- **Бесповторной** называют выборку, при которой отобранный объект в генеральную совокупность не возвращается
- Выборка должна правильно представлять пропорции генеральной совокупности, т.е. она должна быть **представительной (репрезентативной)**
- В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществить случайно: каждый объект выборки отобран случайно из генеральной совокупности, если все объекты имеют одинаковую вероятность попасть в выборку



# Статистическое распределение выборки

---

*Статистическим распределением  
выборки* называют перечень вариантов  
и соответствующих им частот или  
относительных частот



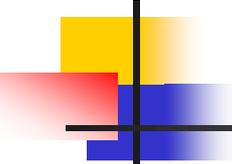
# Пример

---

Число поливов хлопчатника в 2019 г. у 48 фермерских хозяйств.

Распределение частот

<b>X</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>n</b>	<b>2</b>	<b>10</b>	<b>13</b>	<b>19</b>	<b>4</b>



# Пример

---

Число поливов хлопчатника в 2019 г.  
Распределение относительных частот

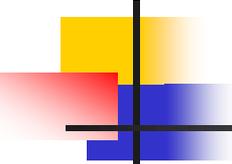
<b>X</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>w</b>	<b>0.04</b>	<b>0.21</b>	<b>0.27</b>	<b>0.4</b>	<b>0.08</b>

# Статистические оценки параметров распределения

- **Несмещенной** называют статистическую оценку, математическое ожидание которой равно оцениваемому параметру при любом объеме выборки, т.е.

$$M(X) = X$$

- **Смещенной** называют оценку, математическое ожидание которой не равно оцениваемому параметру
- **Эффективной** называют статистическую оценку, которая (при заданном объеме выборки) имеет наименьшую возможную дисперсию
- **Состоятельной** называют статистическую оценку, которая при бесконечном увеличении объема выборки стремится по вероятности к оцениваемому параметру. Например, если дисперсия несмещенной оценки при бесконечном увеличении объема выборки стремится к нулю, то такая оценка является и состоятельной



# Генеральная средняя (Population mean)

---

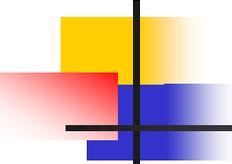
**Генеральной средней** называют среднее арифметическое значений признака генеральной совокупности:

$$pop(\bar{x}) = (x_1 N_1 + x_2 N_2 + \dots + x_k N_k) / N$$

# Выборочная средняя (Sample mean)

**Выборочной средней** называют среднее арифметическое значений признака выборочной совокупности:

$$\text{sample}(\bar{x}) = (x_1 n_1 + x_2 n_2 + \dots + x_k n_k) / n$$



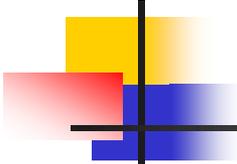
# Пример

---

Найти выборочную среднюю числа поливов

X	2	3	4	5	6
n	2	10	13	19	4

$$\text{sample}(\bar{x}) = (2 \cdot 2 + 3 \cdot 10 + 4 \cdot 13 + 5 \cdot 19 + 6 \cdot 4) / 48 = 205 / 48 = 4,3$$



# Задание

---

Найти выборочную среднюю

X	5	7	4	1
P	3	8	9	2

# Генеральная дисперсия (population variance)

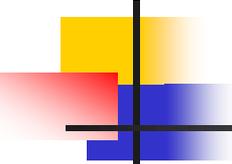
**Генеральной дисперсией** называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения:

$$pop(s^2) = (\sum (x_i - \bar{x})^2) / N$$

# Выборочная дисперсия (sample variance)

**Выборочной дисперсией** называют среднее арифметическое квадратов отклонений наблюдаемых значений признака от их среднего значения:

$$sample( s^2 ) = ( \sum (x_i - \bar{x})^2 ) / n$$



# Пример

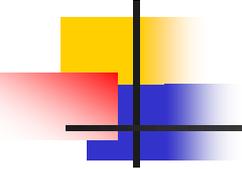
---

Найти выборочную дисперсию числа поливов

X	2	3	4	5	6
n	2	10	13	19	4

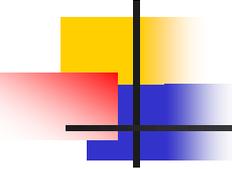
$$sample(s^2) = ((2 - 4,3)^2 + (3 - 4,3)^2 + \dots + (6 - 4,3)^2) / 48 = 10,37 / 48 = 0,22$$

# Свойства выборочной дисперсии



---

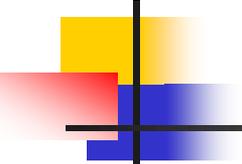
1. Дисперсия постоянной равна нулю.
2. Если ко всем вариантам добавить постоянное число, то дисперсия не изменится.
3. Если все варианты умножить на одно и то же число  $k$ , то дисперсия умножится на  $k^2$



# Задание

---

- Скачать файл «Практическое занятие 4» из MOODLE и выполнить приведенные задания
- Решение задания загрузить в MOODLE



---

СПАСИБО ЗА ВНИМАНИЕ!



+ 998 71 237 1948



[smirzaev@tiame.uz](mailto:smirzaev@tiame.uz)