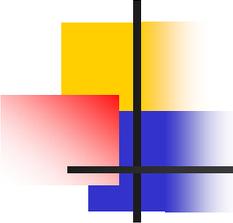
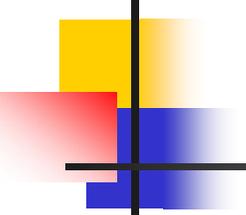


Основные понятия статистики



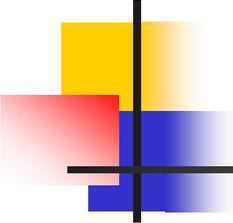
- **Задачи статистики:**
- Указать способы сбора и группировки статистических данных, полученных в результате наблюдений или экспериментов.
- Анализ статистических данных в зависимости от целей исследования.

- **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.



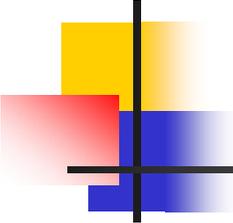
Статистическая обработка данных осуществляется по следующим этапам:

- Планирование процесса сбора данных.
- Подготовка данных для анализа.
- Анализ данных.
- Формулировка выводов на основе анализа данных



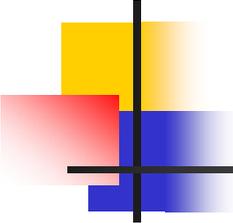
Терминология

- **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**).



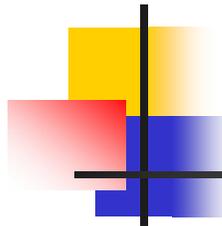
Выборочный метод

- **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

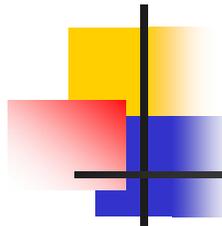
Статистическим распределением выборки называют перечень вариантов и соответствующих им частот или относительных частот.



Пример

- Число поливов хлопчатника в 2004 г.
Распределение частот

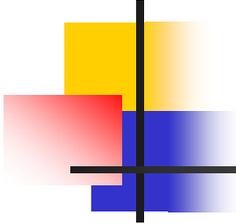
X	2	3	4	5	6
n	2	10	13	19	4



Пример

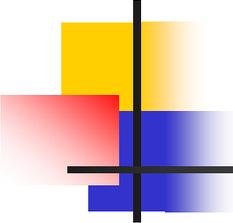
- Число поливов хлопчатника в 2004 г.
Распределение относительных частот

x	2	3	4	5	6
w	0.04	0.21	0.27	0.4	0.08



Статистические оценки параметров распределения

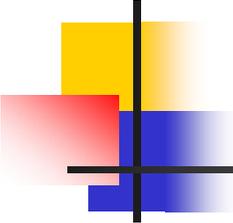
- **Несмещенной** называют статистическую оценку, математическое ожидание которой равно оцениваемому параметру при любом объеме выборки, т.е.
$$M(O) = O.$$
- **Смещенной** называют оценку, математическое ожидание которой не равно оцениваемому параметру.
- **Эффективной** называют статистическую оценку, 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:

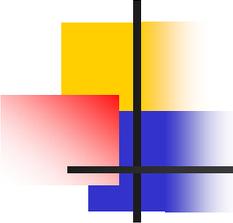
$$\text{pop}(\bar{x}) = (x_1 N_1 + x_2 N_2 + \dots + x_k N_k) / N$$



Выборочная средняя (Sample mean)

Выборочной средней называют среднее арифметическое значений признака выборочной совокупности:

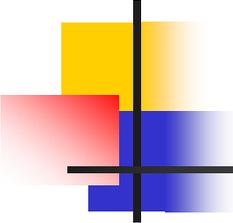
$$\text{sample}(\bar{x}) = (x_1 n_1 + x_2 n_2 + \dots + x_k n_k) / n$$



Генеральная дисперсия (population variance)

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

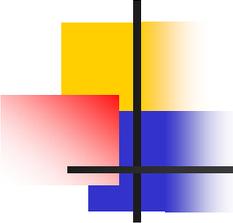
$$pop (s^2) = (\sum (x_i - \bar{x})^2) / N$$



Выборочная дисперсия (sample variance)

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

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



Задания:

- **Readings:**

R.L.Ott, M.Longnecker. An Introduction to Statistical Methods and Data Analysis. Chapters 1,2.

- **K.M.Portier Presentations.** Unit 1 Section 1 (U1S1.ppt)

- **Supplementary Exercices.**

R.L.Ott, M.Longnecker. An Introduction to Statistical Methods and Data Analysis. Chapter 1, pp.14-15.

Deadline: January 27, 2006.