Статистическое управление

Sorochkina O.Y.

Лекция Описательная статистика

Донской государственный технический университет

Кафедра «Управление качеством»

21.03.2020

21 марта 2020 г.

Описательная статистика

Статистические величины и показатели вариации. Абсолютные и относительные статистические величины. Понятие о средних величинах. Виды средних и способы их вычисления. Показатели вариации.

Содержание

1. Статистические показатели.

1.1 Понятие, формы выражения и виды статистических показателей.

Статистический показатель - представляет собой количественную характеристику явлений и процессов в условиях качественной определенности. Качественная определенность показателя заключается в том, что он непосредственно связан с внутренним содержанием изучаемого явления или процесса, его сущностью.

Все статистические показатели *по охвату единиц совокупности* разделяются на: индивидуальные; сводные, а по *форме выражения* на: абсолютные; относительные; средние.

Индивидуальные показатели - характеризуют отдельный объект или отдельную совокупность - предприятие, банк и т.д.

На основе соотнесения 2-х индивидуальных абсолютных показателей, характеризующих один и тот же объект или единицу, получают индивидуальный относительный показатель.

Сводные показатели характеризуют группу единиц, представляющую собой часть статистической совокупности или всю совокупность в целом. Сводные показатели подразделяются на: объемные и расчетные.

- *Объемные показатели* получают путем сложения значений признака отдельных единиц совокупности. Различают абсолютные, относительные и средние объемные показатели.
- *Расчетные показатели* - вычисляются по различным формулам и служат для решения отдельных статистических задач анализа - измерения вариации, характеристики структурных сдвигов и т.д. Расчет. показатели также подразделяются на абсолютные, относительные и средние.

В зависимости от временного фактора различают моментные и интервальные показатели.

Статистические показатели, характеризующие явления и процессы по состоянию *на определенный момент времени* (на определенную дату, начало и конец месяца, года) называются *моментными.*

Статистические показатели, характеризующие явления и процессы *за определенный период* - *день, неделю, месяц, квартал, год* - называются *интервальными.*

1.2. Абсолютные показатели.

Исходной, первичной формой выражения статистических показателей являются *абсолютные величины.* Статистические показатели в форме абсолютных величин характеризуют абсолютные размеры изучаемых статистических процессов и явлений, их массу, площадь, объем и т.д.

Абсолютные статистические показатели всегда являются именованными числами. Они выражаются в натуральных, стоимостных и трудовых единицах измерения

Условно-натуральные показатели используются, когда какой-либо продукт имеет несколько разновидностей, и общий объем можно определить исходя из общего для всех разновидностей потребительского свойства.

1.3. Относительные показатели.

Относительный показатель представляет собой результат деления одного абсолютного показателя на другой и выражает соотношение между количественными характеристиками процессов и явлений.

При расчете относительного показателя абсолютный показатель, находящийся в числителе, называется *текущим (сравниваемым),* а показатель, который находится в знаменателе - *основанием* или *базой.*

Относительные показатели могут выражаться в коэффициентах, процентах или быть именованными.

Относительные показатели можно разделить на следующие виды: динамики, плана, структуры, координации, интенсивности развития, сравнения.

1) *Относительный показатель динамики (ОПД)*- это отношение уровня исследуемого процесса или явления за данный период времени (по состоянию на данный момент времени) к уровню этого же процесса или явления в прошлом.

2) Относительный показатель плана (ОПП) используется предприятиями с целью перспективного планирования своей деятельности.

3) *Относительный показатель структуры (ОПС)* представляет собой соотношение структурных частей изучаемого объекта и их целого.

4) *Относительный показатель координации (ОПК) —* характеризует соотношение отдельных частей целого между собой.

5) Относительный показатель интенсивности (ОПИ) - характеризует степень распространения изучаемого процесса или явления в присущей ему среде:

6) *Относительный показатель сравнения (ОПСр)* - представляет собой соотношение одноименных абсолютных показателей, характеризующих разные объекты (предприятия, фирмы, районы, области и т.п.)

1.4.Графики в статистике

Статистический график - это чертеж, на котором статистические совокупности, характеризуемые определенными показателями, описываются с помощью условных геометрических образов или знаков.

Под *основой графика (графический образ)* понимают геометрические знаки, т.е. совокупность точек, линий фигур, с помощью которых изображаются статистические показатели.

Поле графика - это часть плоскости, где расположены графические образы. Поле графика имеет определенные размеры, которые зависят от его назначения.

Пространственные ориентиры графика задаются в виде системы координат. Система координат необходима для размещения геометрических знаков в поле графика. Наиболее распространенной является система прямоугольных координат.

Масштабные ориентиры статистического графика определяются масштабом и системой масштабных шкал.

Различают *шкалы прямолинейные* (миллиметровая линейка) и *криволинейные* - дуговые и круговые (циферблат часов).

Различают также шкалы равномерные и неравномерные. Шкала является *равномерной,* если равным графическим отрезкам соответствуют равные числовые величины. *Неравномерным шкалам* соответствуют неравные числовые значения.

Плюсы графического изображения

1. наглядно, обозримо, выразительно.
2. сразу видны пределы изменения показателя, сравнительная скорость изменения и колеблемость

Минусы графического изображения

1. Включают меньшее количество данных, чем в таблице.
2. на графике показываются округленные данные, общая ситуация, но не детали.

Статистические графики

Диаграммы	Плоскостные	Столбиковые
Картограммы	Объемные	Ленточные
Картодиаграммы	Полигон	Треугольные
Фоновые	Кумулята	Фигурные
Точечные	Огива	
Линейные	Радиальные	

Классификация и область применения статистических показателей

<table>
<tr><td colspan="4">Статистические показатели</td></tr>
<tr><td rowspan="4">Описательная статистика</td><td rowspan="3">Непрерывные данные</td><td>Коэффициент сдвига</td><td>Среднее (Арифметическое, Геометрическое, Гармоническое) · Медиана · Мода · Размах</td></tr>
<tr><td>Вариация</td><td>Ранг · Среднеквадратическое отклонение · Коэффициент вариации · Квантиль (Дециль, Процентиль/Перцентиль/Центиль)</td></tr>
<tr><td>Моменты</td><td>Математическое ожидание · Дисперсия · Асимметрия · Эксцесс</td></tr>
<tr><td>Дискретные данные</td><td colspan="2">Частота · Таблица контингентности</td></tr>
<tr><td rowspan="6">Статистический вывод и проверка гипотез</td><td>Статистический вывод</td><td colspan="2">Доверительный интервал (Частотная вероятность) · Достоверный интервал (Байесовский вывод) · Статистическая значимость · Мета-анализ</td></tr>
<tr><td>Планирование эксперимента</td><td colspan="2">Генеральная совокупность · Планирование выборки · Районированная выборка · Репликация · Группировка ·Чувствительность и специфичность</td></tr>
<tr><td>Объём выборки</td><td colspan="2">Статистическая мощность · Мера эффекта · Стандартная ошибка</td></tr>
<tr><td>Общая оценка</td><td colspan="2">Байесовская оценка решения · Метод максимального правдоподобия · Метод моментов нахождения оценок · Оценка минимального расстояния · Оценка максимального интервала</td></tr>
<tr><td>Статистические критерии</td><td colspan="2">Z-тест · t-критерий Стьюдента · Критерий Фишера · Критерий Пирсона (Хи-квадрат) · Критерий согласия Колмогорова · Тест Вальда · U-критерий Манна — Уитни · Критерий Уилкоксона · Критерий Краскела — Уоллиса · Критерий Кохрена · Критерий Лиллиефорса</td></tr>
<tr><td>Анализ выживания</td><td colspan="2">Функция выживания · Оценка Каплана — Мейера · Логранк-тест · Интенсивность отказов · Пропорциональная модель опасностей</td></tr>
<tr><td rowspan="3">Корреляция и регрессия</td><td>Корреляция</td><td colspan="2">Коэффициент корреляции Пирсона · Ранг корреляций (Коэффициент Спирмана для ранга корреляций, Коэффициент тау Кендалла для ранга корреляций) · Переменная смешивания</td></tr>
<tr><td>Линейные модели</td><td colspan="2">Основная линейная модель · Обобщённая линейная модель · Анализ вариаций · Ковариационный анализ</td></tr>
<tr><td>Регрессия</td><td colspan="2">Линейная · Нелинейная · Непараметрическая регрессия · Полупараметрическая регрессия · Логистическая регрессия</td></tr>
<tr><td>Графические методы</td><td colspan="3">Столбчатая диаграмма · Совмещённая диаграмма · Диаграмма управления · Лесная диаграмма · Гистограмма · Q-Q диаграмма · Диаграмма выполнения · Диаграмма разброса · Скрипичная диаграмма · Стебель-листья · Ящик с усами</td></tr>
</table>

2. Основные статистические показатели

Основные статистические показатели можно разделить на две группы: меры среднего уровня и меры рассеяния.

Меры среднего уровня

Меры среднего уровня дают усредненную характеристику совокупности объектов по определенному признаку.

- Среднее значение
- Стандартная ошибка
- Стандартное отклонение
- Эксцесс
- Асимметрия
- Интервал
- Минимум
- Максимум
- Счёт
- Медиана
- Мода
- Квантиль
- Математическое ожидание
- Доверительный интервал

Меры рассеяния

Меры рассеяния показывают, насколько хорошо данные значения представляют данную совокупность.

- Дисперсия случайной величины
- Среднеквадратическое отклонение
- Размах вариации
- Интерквантильный размах
- Среднее абсолютное отклонение

3. Средние величины.

3.1. Сущность и значение средних.

Наиболее распространенной формой статистических показателей, используемой в исследованиях, является средняя величина.

Средняя величина представляет собой обобщенную количественную характеристику статистической совокупности в конкретных условиях места и времени.

Сущность средней состоит в том, что она отражает типичный уровень признака и абстрагируется от индивидуальных особенностей, присущих отдельным единицам. Средняя величина только тогда будет отражать типичный уровень признака, когда она рассчитана по качественно однородной совокупности.

Определить среднюю во многих случаях можно через *исходное соотношение средней* (ИСС) или ее логическую формулу:

$$\text{ИСС} = \frac{\text{Суммарное значение или объем осредняемого признака}}{\text{Число единиц или объем совокупности}}$$

В каждом конкретном случае для реализации исходного соотношения требуется одна из форм средней величины. Все виды средних объединяются в общей формуле средней степенной (при различной величине *к*):

1) простая $$\bar{x} = \sqrt[k]{\frac{\sum x_i^k}{n}}$$

2) взвешенная $$\bar{x} = \sqrt[k]{\frac{\sum x_i^k f_i}{\sum f_i}}$$

где k - показатель степени, определяющий вид средней величины;

$\bar{x}$ -средняя величина исследуемого явления;

x_i - i-ый вариант осредняемого признака (i = 1, n)

f_i - вес i-го варианта.

В зависимости от k различают следующие виды ср. величин:

$k = -1$ - средняя гармоническая;

$k = 0$ - средняя геометрическая;

$k = 1$ - средняя арифметическая;

$k = 2$ - средняя квадратическая.

3.2. Средняя арифметическая и ее свойства.

Наиболее распространенным видом средних величин является *средняя арифметическая,* которая в зависимости от характера имеющихся данных может быть простой или взвешенной.

Средняя арифметическая простая применяется, когда значение вариантов встречается по одному числу раз.

$$\bar{x} = \frac{\sum x_i}{n}$$

Средняя арифметическая взвешенная применяется, когда отдельное значение признака повторяется неодинаковое количество раз, т.е. она используется в расчетах средней по сгруппированным данным или вариационным рядам, которые могут быть дискретными и интервальными.

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

При расчете средней по интервальному вариационному ряду для выполнения необходимых вычислений переходят от интервалов к их серединам.

Свойства средней арифметической

1. Произведение средней на сумму частот равно сумме произведений отдельных вариантов на соответствующие им частоты:

$$\bar{x}\sum f_i = \sum x_i f_i$$

2. *Свойство для отклонений:* сумма отклонений вариант от средней арифметической равно нулю:

$$\sum (x_i - \bar{x}) f_i = 0 \text{ или } \sum (x_i - \bar{x}) = 0$$

3. *Свойство для вариант:* если все осредняемые уменьшить или увеличить на постоянное число А, то средняя арифметическая соответственно уменьшится или увеличится на ту же

$$\bar{x} = \frac{\sum (x_i \pm A) f_i}{\sum f_i} = \bar{x} \pm A$$

величину:

4. Если варианту увеличить или уменьшить в какое-то число раз, то в то же число раз увеличится или уменьшится среднее арифметическое:

$$\frac{\sum(x_i d) f_i}{\sum f_i} = \bar{x} d \qquad \text{или} \qquad \frac{\sum\left(\frac{x_i}{d_i}\right) f_i}{\sum f_i} = \frac{\bar{x}}{d}$$

5. *Свойство для частот:* если частоты (веса) ряда увеличить или уменьшить на произвольное число, то средняя арифметическая от этого не изменится:

$$\frac{\sum x_i (f_i / A)}{\sum (f_i / A)} = \bar{x}$$

6. Если веса или частоты всех вариант равны между собой, то средняя арифметическая взвешенная будет равна средней арифметической простой:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum x_i}{n} \quad \text{если} \quad f_i = f_1$$

Знание основных свойств средней арифметической позволяет упростить ее вычисление особенно для вариационного ряда с равными интервалами, т.е. способом моментов:

$$\bar{x}_M = \frac{\sum\left(\frac{x - A}{i}\right) f}{\sum f} \cdot i + A$$

где i- интервал,

x - серединное значение интервала,

A - условная величина,

f - частота признака. За (A) условную величину принимают варианту, занимающую серединное положение в данном ряду и имеющую наибольшую частоту.

<u>*Доминирующее серединное положение в ряду:*</u>

$$\bar{x} = m_1 \cdot i + A$$

$$m_1 = \left(\sum\left(\frac{x - A}{i}\right) f\right) / \sum f$$

Серединное m_iиз значений $\frac{(x-A)}{i}$ называется <u>*моментом первого порядка.*</u>

3.3. Другие виды средних.

<u>*Средняя гармоническая*</u> ~ это величина, обратная средней арифметической, когда κ = -1. Когда статистическая информация не содержит частот по отдельным вариантам совокупности, а представлена как их произведение, применяется формула <u>*средней гармонической взвешенной:*</u>

$$\bar{x} = \frac{\sum w_i}{\sum w_i / x_i} \quad , \quad \text{где } w_i = x_i \cdot f_i$$

Когда объемы явлений, т.е. произведения (w, = w,). по каждому признаку равны, применяется <u>*средняя гармоническая простая*</u>

$$\bar{x} = \frac{n}{\sum 1/x}$$

где $\sum 1/x$ – сумма обратных значений вариант,
n – число вариант.

Средняя геометрическая - это величина, используемая как средняя из отношений или в рядах распределения, представленных в виде геометрической прогрессии, когда *к = О,*

$$\bar{x} = \sqrt[n]{\prod(x)}.$$

Средняя геометрическая используется в расчетах среднегодовых темпов роста и для определения равноудаленной величины от минимального и максимального значений признака.

средняя геометрическая простая:

$$\bar{x} = \sqrt[n]{\prod(x_i)}$$

средняя геометрическая взвешенная:

$$\bar{x} = \sqrt[\sum f]{\prod(x_i)^{f_i}}$$

Средние диаметры колес, труб, средние стороны квадратов определяются при помощи *средней квадратической.*

Среднеквадратические величины используются для расчета некоторых показателей, например коэффициент вариации, характеризующего ритмичность выпуска продукции. Здесь определяют среднеквадратическое отклонение от планового выпуска продукции за определенный период по следующей формуле:

$$X_{kvad} = \sqrt{\sum \Delta x * x / N}$$

Эти величины точно характеризуют изменение экономических показателей по сравнению с их базисной величиной, взятое в его усредненной величине.

Квадратическая простая

Средняя квадратическая простая вычисляется по формуле:

$$\bar{x}_{квадр} = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}},$$

$$\text{или } \overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n}.$$

Квадратическая взвешенная

Средняя квадратическая взвешенная равна:

$$\bar{x}_{квадр} = \sqrt{\frac{x_1^2 f_1 + x_2^2 f_2 + \ldots + x_n^2 f_n}{f_1 + f_2 + \ldots + f_n}} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}},$$

$$\text{или } \overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2 f_i}{\sum_{i=1}^{n} f_i}.$$

Формула степенной простой в общем виде

$$\bar{x} = \left(\frac{\sum_{i=1}^{n} x_i^k}{n}\right)^{\frac{1}{k}} = \sqrt[k]{\frac{\sum_{i=1}^{n} x_i^k}{n}} = \sqrt[k]{\frac{x_1^k + x_2^k + \ldots + x_n^k}{n}},$$

где:

- x_i — индивидуальное значение признака i-й единицы совокупности
- k — показатель степени средней величины
- n — число единиц совокупности

Формула степенной средней взвешенной в общем виде

$$\bar{x} = \left(\frac{\sum_{i=1}^{n} x_i^k f_i}{\sum_{i=1}^{n} f_i}\right)^{\frac{1}{k}} = \sqrt[k]{\frac{\sum_{i=1}^{n} x_i^k f_i}{\sum_{i=1}^{n} f_i}} = \sqrt[k]{\frac{x_1^k f_1 + x_2^k f_2 + \ldots + x_n^k f_n}{f_1 + f_2 + \ldots + f_n}},$$

где:

- f_i — частота повторения i-й варианты.

В зависимости от того, какое значение принимает показатель степени средней величины k, получаем различные виды средних:

Вид степенной средней	Показатель степени средней (k)	Формула расчета	
		Простая	Взвешенная
Гармоническая	- 1	$\bar{x} = \dfrac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$	$\bar{x} = \dfrac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} \frac{x_i f_i}{x_i}}$
Геометрическая	1	$x = \sqrt[n]{x_1 x_2 \dots x_n}$	$\bar{x} = \sqrt[\sum f_i]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}}$
Арифметическая	0	$\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$	$\bar{x} = \dfrac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$
Квадратическая	2	$\bar{x} = \sqrt{\dfrac{\sum_{i=1}^{n} x_i^2}{n}}$	$\bar{x} = \sqrt{\dfrac{\sum_{i=1}^{n} x_i^2 f_i}{\sum_{i=1}^{n} f_i}}$

При расчете различных степенных средних по одним и тем же данным значения средних будут неодинаковыми. Чем выше показатель степени (k), тем больше величина средней, т.е. действует правило мажорантности средних:

$$\bar{x}_{гарм} \leq \bar{x}_{геом} \leq \bar{x}_{арифм} \leq \bar{x}_{квадр} \leq \bar{x}_{куб} \,.$$

3.4. Структурные средние величины.

Для характеристики структуры совокупности применяются особые показатели - *структурные средние*. К таким показателям относятся мода и медиана.

Модой (Мо) называется чаще всего встречающийся вариант.

В *дискретном ряду* мода - это варианта с наибольшей частотой. В *интервальном ряду* модой считают центральный вариант модального интервала, т.е. того интервала, который имеет наибольшую частоту (частотность).

Мода для интервального ряда:

$$M_o = x_{M_o} + i_{M_o} \frac{(f_{M_o} - f_{M_{o-1}})}{(f_{M_o} - f_{M_{o-1}}) + (f_{M_o} - f_{M_{o+1}})}$$

где $x_м$ - нижняя граница модального интервала,

i_{Mo} - величина модального интервала,

f_{Mo} - частота, соответствующая модальному интервалу.

f_{Mo-1} - частота, предшествующая модальному интервалу,

f_{Mo+1}- частота интервала, следующего за модальным.

Медиана (Ме) - это величина, которая делит численность упорядоченного вариационного ряда на две равные части: одна часть имеет значение варьирующего признака меньше, чем средний вариант, а другая - больше.

Для ранжированного ряда (т.е. построенного в порядке возрастания или убывания индивидуальных величин) *с нечетным числом,* членов медианой является варианта, расположенная в центре ряда.

Для ранжированного ряда с *четным числом членов* медианой будет средняя арифметическая из двух смежных вариант.

$$M_e = \left(x_{M_e} + x_{M_{e+1}}\right)/2$$

Для интервального вариационного ряда:

$$M_e = x_{M_e} + i_{M_e} \frac{\frac{\sum f}{2} - \sum f_{M_{e-1}}}{f_{M_e}}$$

где x_{M_e} - нижняя граница медианного интервала;

i_{M_e} - величина медианного интервала;

$\frac{\sum f}{2}$ - полусумма частот ряда;

$\sum f_{M_{e-1}}$ - сумма накопленных частот, предшествующих медианному интервалу,

f_{M_e} - частота медианного интервала.

Медианный интервал - это интервал, где сумма накопленных частот составляет половину (или больше) всей суммы частот ряда.

f	f'
10	10
20	30(10+20)
30	60(30+30)

4. Показатели вариации

4.1. Понятие вариации и ее значение.

Термин «вариация» произошел от латинского varito –изменение, колеблемость, различие. Однако не всякое различие называется вариацией. Под *вариацией* в статистике понимают такие количественные изменения величины исследуемого признака в пределах однородной совокупности, которые обусловлены перекрещивающимся влиянием действия различных факторов.

Исследование вариации в статистике имеет важное значение, т.к. дает возможность оценить степень воздействия на данный признак других варьирующих признаков. Определение вариации необходимо при организации выборочного наблюдения, построения статистических моделей, разработке материалов экспертных опросов и т.д.

Средняя величина – это обобщающая характеристика признака изучаемой совокупности. Она не дает представления о том, как отдельные значения изучаемого признака группируются вокруг средней. Поэтому для характеристики колеблемости признака используют показатели вариации.

4.2. Показатели вариации и их значение в статистике.

Показатели вариации делятся на две группы: абсолютные и относительные.

К ***абсолютным*** относятся: размах вариации, среднее линейное отклонение, дисперсия и среднее квадратичное отклонение.

1. Самым распространенным абсолютным показателем является *размах вариации,* определяемый как разность между наибольшим (Xmax) и наименьшим (Xmin) значениями вариантов.

$$R = x_{\max} - x_{\min}$$

Этот показатель прост для расчета, что и обусловило его широкое распространение. Однако, он улавливает только крайние отклонения и не отражает отклонений всех вариант в ряду.

2. Для обобщающей характеристики распределения отклонений рассчитывают *среднее линейное отклонение* $\bar{d}$, определяемое как средняя арифметическая из отклонений индивидуальных значений от средней, без учета знака этих отклонений:

$\bar{d} = \dfrac{\sum |x_i - \bar{x}|}{n}$ - невзвешенное среднее линейное отклонение

$\bar{d} = \dfrac{\sum |x_i - \bar{x}| f_i}{\sum f_i}$ - взвешенное среднее линейное отклонение

Среднее линейное отклонение как меру вариации признака применяют в статистической практике редко, т.к. во многих случаях этот показатель не устанавливает степень рассеивания.

3. Меру вариации более объективно отражает показатель *дисперсии* (σ^2- средний квадрат отклонений), определяемый как средняя из отклонений, возведенных в квадрат:

$\sigma^2 = \dfrac{\sum (x_i - \bar{x})^2}{n}$ - невзвешенная или $\sigma^2 = \dfrac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$ - взвешенная

4. Корень квадратный из дисперсии σ «среднего квадрата отклонений» представляет собой *среднее квадратическое отклонение:*

$\sigma = \sqrt{\sigma^2}$

Среднее квадратическое отклонение (СКО) выражается в тех же единицах измерения, что и признак (в литрах, тоннах, рублях, %-х и т.д.). СКО является мерилом надежности средней. Чем меньше СКО, тем лучше средняя арифметическая отражает собой представляющую совокупность.

К **относительным показателям**, позволяющим сравнивать характер рассеивания в различных распределениях, относятся следующие:

1. *Коэффициент осцилляции* – отражающий относительную колеблемость крайних значений признака вокруг средней:

$$V_R = \frac{R}{\bar{x}} \cdot 100\%$$

2. *Относительное линейное отклонение* характеризует долю усредненного значения абсолютных отклонений от средней величины:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} 100\%$$

3. *Коэффициент вариации* является наиболее распространенным показателем колеблемости, используемым для оценки типичности средней величины.

$$V_\sigma = \frac{\sigma}{\bar{x}} 100\%$$

Если ν>33% , то это говорит о большой колеблемости признака в изучаемой совокупности.

4.3. Дисперсия: свойства и методы расчета.

Дисперсия обладает рядом свойств, которые позволяют упростить ее расчеты.

1). Если из всех значений вариант отнять какое-то постоянное число А, то средний квадрат отклонений от этого не изменится:

$$\sigma^2(x_i - A) = \sigma^2$$

2) Если все значения вариант разделить на какое-то постоянное число А, то средний квадрат отклонений уменьшится от этого в A^2 раз, а среднее квадратическое отклонение – в А раз.

$$\sigma^2\left(\frac{x_i}{A}\right) = \sigma^2 : A^2$$

3). Если исчислить средний квадрат отклонений от любой величины А, которая в той или иной степени отличается от средней арифметической х, то он всегда будет больше среднего квадрата отклонений σ^2, исчисленный от средней арифметической:

$$\sigma_A^2 \rangle \sigma_x^2$$

А именно средний квадрат отклонений при этом будет больше на квадрат разности средней и этой условно взятой величиной, т.е. на $(\bar{x} - A)^2$:

$$\sigma_A^2 = \sigma_{\bar{x}}^2 + (\bar{x} - A)^2$$

или

$$\sigma_{\bar{x}}^2 = \sigma_A^2 - (\bar{x} - A)^2$$

Дисперсия от средней имеет *<u>свойство минимальности</u>*, т.е. она всегда меньше дисперсий, исчисленных от любых других величин. В этом случае, когда А приравнивается к нулю формула принимает вид:

$$\sigma^2 = \frac{\sum x^2 f}{\sum f} - \left(\frac{\sum xf}{\sum f}\right)^2 \quad \text{или} \quad \sigma^2 = \overline{x}^2 - (\bar{x})^2$$

средний квадрат значений признака	квадрат среднего значения признака
m_2	m_1^2

Этот способ расчета дисперсии и среднего квадратического отклонения называется *<u>способом моментов,</u>* или способом от условного нуля. Он применяется при условии равных интервалов.

Используя второе свойство дисперсии, разделив, все варианты на величину интервала, получим: $\sigma^2 = i^2(m_2 - m_1^2)$.

4.4. Виды дисперсий и закон сложения дисперсий.

Наряду с изучением вариации признака по всей совокупности в целом часто бывает необходимо проследить количественные изменения признака по группам, на которые разделяется совокупность, а также между группами. Такое изучение вариации достигается посредством вычисления и анализа различных видов дисперсии.

Выделяют общую, межгрупповую и внутригрупповую дисперсии.

<u>Общая дисперсия</u> характеризует вариацию признака, которая зависит от всех условий в данной совокупности:

$$\sigma_o^2 = \frac{\sum (x_i - \bar{x}_o)^2 f_i}{\sum f_i}$$

где $\bar{x}_o$ - общая средняя для всей изучаемой совокупности

Межгрупповая дисперсия отражает вариацию изучаемого признака, которая возникает под влиянием признака-фактора, положенного в основу группировки. Она характеризует колеблемость групповых (частных) средних $\bar{x}_i$ около общей средней $\overline{x_o}$:

$$\sigma^2 = \frac{\sum(\bar{x}_i - \bar{x}_o)^2 f_i}{\sum f_i}$$

где $\bar{x}_i$ – средняя по отдельным группам,

$\bar{x}_o$ – общая средняя,

f_i - численность отдельных групп.

Средняя внутригрупповая дисперсия характеризует случайную вариацию в каждой группе. Эта вариация возникает под влиянием других не учитываемых факторов и не зависит от условия (признака-фактора), положенного в основу группировки:

$$\bar{\sigma}_i^{\,2} = \frac{\sum \sigma_i^2 f_i}{\sum f_i}$$

Существует закон, связывающий три вида дисперсии (правило сложения дисперсий): *общая дисперсия равна сумме средних из внутригрупповой и межгрупповой дисперсии:*

$$\sigma_o^2 = \bar{\sigma}_i^2 + \delta^2$$

5. Список литературы

1. Ефимов В.В. Статистические методы в управлении качеством. Ульяновск: УлГТУ, 2003 – 134 с.
2. Ефимов В.В. Основы бережливого производства : учебное пособие / В. В. Ефимов. – Ульяновск : УлГТУ, 2011. – 160 с.
3. Энциклопедия статистических терминов. Методические основы статистики. Том 1 - М.: Федеральная служба государственной статистики, 2013
4. Статистические методы управления качеством //www.lenobl.ru, 2005.
5. Климанов В. Статистические методы управления качеством//victor61058.narod.ru, 2004.
6. Окрепилов В.В. Управление качеством. СПб.: Наука, 2000. - 911 с.
7. ГОСТ Р 50779.10-2000 Статистические методы. Вероятность и основы статистики. Термины и определения. – М:Стандартинформ, 2005
8. ГОСТ Р 50779.11-2000 Статистические методы. Статистическое управление качеством. Термины и определения– М:Стандартинформ, 2005
9. ГОСТ Р ИСО/ТО 10017-2005 Статистические методы. Руководство по применению в соответствии с ГОСТ Р ИСО 9001– М:Стандартинформ, 2005

ПРИЛОЖЕНИЕ А

Числовые характеристики случайных величин			
N п/п	**Характеристика (параметр) случайной величины и ее обозначение**	**Формула для нахождения характеристики случайной величины**	**Примечание**
1	Математическое ожидание M_x	$M_x = \sum_{i=1}^{n} p_x x_i$	Характеризует положение случайной величины на числовой оси
2	Среднее значение $\overline{x}$	$\overline{x} = \frac{1}{N}\sum_{i=1}^{n} x_i$	Если случайная величина независимая, то $\overline{x} = M_x$
3	Мода M_o	Это такое значение x_i, для которого p_i наибольшее	Равна наиболее часто встречающемуся значению x_i. Если таких значений в вариационном ряду несколько, то M_o не определяется.
4	Медиана M_e	Если n четное, то $M_e = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}$ Если n нечетное, то $M_e = x_{\frac{n+1}{2}}$	Это такое значение, которое находится в центре ранжированного ряда.
5	Дисперсия D_x	$D_x = M(X - M_x)^2 = \sum_{i=1}^{n} p_i(x_i - \overline{x}$	Характеризует рассеяние случайной величины вокруг среднего значения
6	Среднеквадратичное отклонение (стандарт) σ	$\sigma = \sqrt{D_x}$	Характеризует действительное рассеяние случайной величины вокруг среднего значения.
7	Коэффициент вариации V	$V = \frac{\sigma}{X}$	Наряду с дисперсией характеризует изменчивость случайной величины
8	Центрированное нормированное уклонение $\overline{x_i}$	$\overline{x_i} = \frac{(x_i - \overline{x})}{\sigma}$	
9	Начальный момент k-го порядка m_k	$m_k = M_X^k = \sum_{i=1}^{n} p_x x_i^k$	
10	Центральный момент k-го порядка μ_k	$\mu_k = M(X - M_x)^k = \sum_{i=1}^{n} p_i(x_i - \overline{x})^k$	
11	Асимметрия распределения A	$A = \frac{\mu_3}{\sigma^3}$	
12	Эксцесс распределения E	$E = \frac{\mu_4}{\sigma^4} - 3$	

6. Контрольные вопросы.

1. Дайте определение термину статистический показатель.
2. Приведите классификацию статистических показателей.
3. Что такое статистических график?
4. Достоинства и недостатки статистических графиков.
5. Разновидность статистических графиков.
6. Область применения статистических показателей.
7. Какие статистические показатели относятся к основным?
8. Меры среднего уровня (положения)
9. Меры рассеяния.
10. В чем сущность и значения средних?
11. Виды средних.
12. Каковы свойства средней арифметической?
13. Структурные средние величины.
14. Что такое Мода, формула.
15. Что такое Медиана, формула.
16. Приведите формулы средних.
17. Что такое вариация?
18. Классификация вариации.
19. Какие свойства дисперсии вы знаете.
20. Приведите классификацию дисперсий.
21. В чем суть закона сложения дисперсий?