



Таможенная статистика

Лекция 2



ТЕМА 4. СИСТЕМА ПОКАЗАТЕЛЕЙ И ПРИЗНАКОВ В ТАМОЖЕННОЙ СТАТИСТИКЕ

Система показателей и признаков в таможенной статистике

- Ключевым элементом таможенной статистики является *система показателей*, отражающих цифровую характеристику различных экономических явлений и процессов, а также экономики в целом.
- Под термином «система показателей» понимается некоторое упорядоченное множество *взаимосвязанных и взаимосогласованных показателей*, характеризующих экономику страны в целом и основные аспекты внешней торговли в частности.

Определение содержания показателя и методов его оценки принято называть *разработкой методологии*.

Этапы разработки методологии

- 1) идентификация явлений и процессов, подлежащих статистическому изучению (определение типа данных, требующих разработки), формулирование целей, ради которых должны быть исчислены те или иные показатели (например, целью изучения импорта является измерение потребностей населения, их динамики, состояния внешнеэкономических связей страны и т.д.);
- 2) определение содержания показателей (например, при исчислении показателя экспорта должны быть точно определены те виды товаров, которые подлежат и не подлежат включению в этот показатель);

Этапы разработки методологии

- 3) определение методов оценки отдельных показателей, например, типа цен, которые должны быть использованы для оценки товаров при исчислении величины внешнеторгового оборота: цены ФОБ, СИФ и т. д.;
- 4) определение основных классификаций, которые должны быть применены для распределения изучаемых экономических явлений на однородные группы на основе тех или иных критериев (например, ТН ВЭД);
- 5) определение основных источников данных, необходимых для исчисления показателей, а также процедуры обработки собранных данных с целью получения обобщающих

Система показателей таможенной статистики

Система показателей таможенной статистики должна соответствовать определенным требованиям, чтобы была возможность проводить описание и анализ их развития.

Во-первых, она должна иметь *всеохватывающий* характер, т.е. распространяться на все аспекты исследуемого процесса, должны быть охвачены все хозяйствующие субъекты, все виды экономических операций, которые они выполняют.

Во-вторых показатели системы, относящиеся к различным аспектам экономического процесса, должны быть методологически *взаимосогласованы*, т.е. они должны быть основаны на гармонизированных концепциях, определениях и классификациях.

Система показателей и признаков

Система показателей и признаков, разрабатываемых таможенной статистикой внешней торговли на основе сведений ГТД, содержит:

- | | |
|--|--|
| 1) отчетный период; | 10) вес нетто; |
| 2) направление <u>товаропотока</u> (ввоз или вывоз); | 11) код и наименование дополнительных единиц измерения; |
| 3) страна происхождения; | 12) количество товара в дополнительных единицах измерения; |
| 4) страна назначения; | 13) характер сделки; |
| 5) торгующая страна; | 14) таможенный режим; |
| 6) страна отправления; | 15) особенность декларирования товара; |
| 7) статистическая стоимость; | 16) регион. |
| 8) код и наименование товара; | |
| 9) вес брутто; | |

Система показателей и признаков

- Период, к которому относится информация о внешнеторговой операции, определяется в соответствии с моментом учета. В таможенной статистике внешней торговли учет ввоза и вывоза товаров при водных, железнодорожных, автомобильных, воздушных перевозках производится при декларировании товара и ведется по дате выпуска товара, проставленной в ГТД.
- Данная система показателей и признаков информирует заинтересованные стороны о том, что ввозится и вывозится и в каком объеме (как в натуральном, так и в стоимостном выражении), а также какие зарубежные страны выступают партнерами России и каковы масштабы внешнеторговых сделок с ними.

Система показателей и признаков

Особенности декларирования товаров, перемещаемых трубопроводным транспортом (нефть, нефтепродукты, газ, вода и др.) и по линиям электропередачи, обусловлены тем, что их перемещение осуществляется большими партиями в непрерывном режиме в течение длительного времени.

Поэтому статистический учет, в связи с особенностью их декларирования, осуществляется по дате последнего транспортного документа, а для природного газа и электроэнергии – по дате последнего дня месяца, в котором поставлялся товар.

Классификация и кодирование товаров

Для классификации и кодирования товаров в таможенной статистике внешней торговли России применяется классификатор «Товарная номенклатура внешнеэкономической деятельности РФ» (ТН ВЭД России), структура которого представлена в Приложении 5. ТН ВЭД России основана на номенклатуре Гармонизированной системы описания и кодирования товаров (ГС) Всемирной таможенной организации (далее – ВТО). Схема построения ТН ВЭД России совпадает с ГС, которая имеет 5 уровней детализации товаров.

В ТН ВЭД России товары располагаются по степени их обработки (сырье, полуфабрикаты, готовые изделия)

ЕДИНАЯ ТОВАРНАЯ НОМЕНКЛАТУРА ВНЕШНЕЭКОНОМИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ														
Раздел I: ЖИВЫЕ ЖИВОТНЫЕ, ПРОДУКТЫ ЖИВОТНОГО ПРОИСХОЖДЕНИЯ					Раздел II: ПРОДУКТЫ РАСТИТЕЛЬНОГО ПРОИСХОЖДЕНИЯ									
Раздел III			Раздел IV: ГОТОВЫЕ ПИЩЕВЫЕ ПРОДУКТЫ						Раздел V: Минеральные продукты					
Раздел VI: ПРОДУКЦИЯ ХИМИЧЕСКОЙ ПРОМЫШЛЕННОСТИ														
Раздел VII: ПЛАСТМАССА И РЕЗИНА		Раздел VIII: ШКУРЫ, КОЖА, МЕХ И ИЗДЕЛИЯ ИЗ НИХ		Раздел IX: ДРЕВЕСИНА И ИЗДЕЛИЯ			Раздел X: МАССА ИЗ ДРЕВЕСИНЫ, БУМАГА, КАРТОН							
Раздел XI: ТЕКСТИЛЬНЫЕ МАТЕРИАЛЫ И ТЕКСТИЛЬНЫЕ ИЗДЕЛИЯ														
							ДЕЛОВАЯ СОЦИАЛЬНАЯ СЕТЬ WINGI.RU ПОРТАЛ ПО ЛОГИСТИКЕ И ТАМОЖНЕ							
Раздел XII: ОБУВЬ, ГОЛОВНЫЕ УБОРЫ, ЗОНТЫ, ТРОСТИ			Раздел XIII: КАМЕНЬ, КЕРАМИКА, СТЕКЛО											
Раздел XIV:		Раздел XV: НЕДРАГОЦЕННЫЕ МЕТАЛЛЫ И ИЗДЕЛИЯ ИЗ НИХ												
Раздел XVI: МАШИНЫ, ОБОРУДОВАНИЕ		Раздел XVII: ТРАНСПОРТ, ЛЕТАТЕЛЬНЫЕ АППАРАТЫ, ПЛАВУЧИЕ СРЕДСТВА			Раздел XVIII: ИНСТРУМЕНТЫ И АППАРАТЫ		Раздел XIX: ОРУЖИЕ		Раздел XX: РАЗНЫЕ ПРОМЫШЛЕННЫЕ ТОВАРЫ		Раздел XXI			

В ТН ВЭД России товары располагаются по степени их обработки (сырье, полуфабрикаты, готовые изделия)

Первый (высший) уровень предусматривает группировку товаров в Разделы - с I по XXI.

С целью конкретизации товаров, относящихся к соответствующему разделу, используются примечания.

В ТН ВЭД России товары располагаются по степени их обработки (сырье, полуфабрикаты, готовые изделия)

Второй уровень объединяет товары в товарные группы - с 01 по 97, исключая товарную группу 77, зарезервированную Комитетом по ГС ВТО для целей развития номенклатуры.

Группы формируются по таким критериям, как, например:

- материал, из которого изготовлен товар (Товарные группы 39-46, 70-81 и др.);
- функциональное предназначение товара (Товарные группы 30-34, 36, 37, 64-66, 84-97 и др.);
- степень обработки товара (от сырья до товара, прошедшего высокую степень обработки).

Для целей конкретизации товарных групп также используются *Примечания к товарным группам*.

В ТН ВЭД России товары располагаются по степени их обработки (сырье, полуфабрикаты, готовые изделия)

Третий уровень группирует товары по товарным позициям (всего – 1244 позиции).

В товарных позициях товары детализируются с учетом таких признаков как вид товара, его форма. Причем уровень конкретизации товара уже настолько точен, что описание товара имеет юридическое (правовое) значение и, как правило, не требует дополнительных примечаний.

В ТН ВЭД России товары располагаются по степени их обработки (сырье, полуфабрикаты, готовые изделия)

- На четвертом и пятом уровнях товары детализируются в субпозиции и подсубпозиции соответственно. При отсутствии полного текста описывающего товар допускается использование примечаний.

Таким образом, неотъемлемой частью ТН ВЭД являются примечания к разделам, группам, товарным позициям, субпозициям и подсубпозициям, а также *Основные правила интерпретации ТН ВЭД.*

ТН ВЭД России

В соответствии с соглашением о единой Товарной номенклатуре в РФ применяется Товарная номенклатура внешнеэкономической деятельности Содружества Независимых Государств (ТН ВЭД СНГ).

В ТН ВЭД СНГ без каких-либо дополнений и изменений использованы все товарные позиции и субпозиции ГС, относящиеся к ним цифровые коды, а также основные правила интерпретации ГС.

Объем внешней торговли

Объем внешней торговли региона или страны в целом характеризуется системой показателей, которая включает:

- абсолютные величины в натуральном выражении – объем вывезенных и ввезенных товаров по видам;
- абсолютные стоимостные величины: объем экспорта (импорта) – всего, в том числе по видам вывезенных (ввезенных) товаров; объем внешнеторгового оборота; сальдо внешней торговли.

Все эти абсолютные величины являются интервальными, исчисляемыми за определенный период: месяц, квартал, год.

Объем внешней торговли

В масштабах отдельной страны (региона) объем *внешнеторгового оборота (ВО)* складывается из суммы стоимости экспорта (*Э*) и импорта (*И*), то есть по формуле

$$ВО = Э + И$$

Сальдо внешней торговли (торгового баланса) страны (региона) (*СВТ*) определяется как разница между суммами экспорта (*Э*) и импорта (*И*), то есть по формуле:

$$СВТ = Э - И$$

Если СВТ положительно, значит экспорт превышает импорт, то есть торговый баланс активный,

Если СВТ отрицательно, значит импорт больше экспорта, а торговый баланс пассивный.

Если СВТ = 0, то такое соотношение в торговом балансе называется нетто-балансом.

Объем внешней торговли

Сравнение экспорта с импортом может быть и относительным, при этом получается индекс координации, называемый *коэффициентом покрытия импорта экспортом*, определяемый по формуле:

$$K_{\text{покр}} = \frac{\text{Э}}{\text{И}}$$

Помимо стоимостных показателей внешнеторгового оборота и сальдо торгового баланса, определяемых на уровне отдельных стран, в международной статистике внешней торговли исчисляют оборот и сальдо мировой торговли.

Оборот мировой торговли

Оборот мировой торговли (ОМТ) характеризует общий объем перемещаемых между странами товаров и рассчитывается как сумма стоимости экспорта всех стран, то есть по формуле:

$$ОМТ = \sum_{i=1}^n Э_i$$

где n – число стран, осуществивших в отчетном периоде экспорт товаров.

Такая методика расчета ОМТ объясняется тем, что экспорт товаров из всех стран мира соответствует импорту в эти страны, поэтому суммирование объемов мирового экспорта и импорта по аналогии с предыдущей формулой привело бы к двойному счету одних и тех же товаров.

Сальдо мировой торговли

За счет различий в базисных ценах экспорта и импорта (экспорт – в ценах FOB, импорт – в ценах CIF) возникает величина, именуемая в международной статистике как *сальдо мировой торговли (СМТ)*, определяемая как разница между суммами стоимости экспорта и импорта всех стран мира, то есть по формуле

$$СМТ = \sum_{i=1}^n \mathcal{E}_i - \sum_{j=1}^m \mathcal{I}_j$$

- где m – число стран, импортировавших товары в отчетном периоде.

СМТ показывает, во что мировому сообществу обходится доставка товаров до стран-импортеров, при этом всегда $СМТ < 0$, так как цены CIF на перемещаемые между странами товары всегда превышают цены FOB.

Статистика внешней торговли

Статистика внешней торговли изучает участие отдельных стран в международном разделении труда.

Вовлеченность национальных экономик в мирохозяйственные связи отражается в системе показателей, в основе расчета которых лежат относительные статистические величины

Основные статистические величины

доля отдельных стран или групп стран в мировой торговле определяется по формуле:

$$d_j = \frac{\mathcal{E}_j}{OMT} \quad d_n = \frac{\sum_{j=1}^n \mathcal{E}_j}{OMT}$$

где \mathcal{E}_j – общий объем экспорта j -й страны;
 n – число стран анализируемой группы, экспортировавших товары.

Основные статистические величины

Доля отдельных стран в экспорте отдельных товаров (товарных групп) показывает в рамках каких отраслей и видов производств развивается специализация страны в международном разделении труда, определяется по формуле:

$$d_{ij} = \frac{\mathcal{E}_{ij}}{\sum_{j=1}^m \mathcal{E}_{ij}}$$

где \mathcal{E}_{ij} – объем экспорта i -го товара j -й страны;

m – число стран, экспортировавших i -й товар на мировой рынок.

Доля экспорта

доля экспорта отдельной страны в валовом национальном продукте (ВНП) показывает, какую часть произведенного ВНП страна направляет на внешний рынок или сколько на единицу ВНП приходится единиц вывезенных товаров, определяется по формуле

$$d_j = \frac{\mathcal{E}_j}{\text{ВНП}_j}$$

где ВНП_j – объем валового национального продукта j -й страны.

Коэффициент зависимости

коэффициент зависимости национальной экономики от импорта показывает, сколько на единицу произведенного ВВП приходится единиц ввезенных товаров, определяется по формуле:

$$K_{зав j} = \frac{I_j}{ВВП_j}$$

где I_j – общий объем импорта j -й страны.

Доля экспорта

*доля экспорта в производстве
отдельных видов продукции
определяется по формуле:*

$$d_{\text{э}j} = \frac{\text{Э}_{ij}}{Q_{ij}}$$

где Q_{ij} – объем производства i -го товара
 j -й страны.

Коэффициент зависимости

коэффициент зависимости национальной экономики от импорта показывает, сколько на единицу произведенного ВВП приходится единиц ввезенных товаров, определяется по формуле

$$K_{зав j} = \frac{I_j}{ВВП}$$

где I_j – общий объем импорта j -й страны

Доля экспорта

*доля экспорта в производстве
отдельных видов продукции
определяется по формуле*

$$d_{\text{Э}j} = \frac{\text{Э}_{ij}}{Q_{ij}}$$

где Q_{ij} – объем производства i -го товара
 j -й страны.

Доля импорта

Доля импорта в потреблении отдельных видов продукции показывает зависимость экономики страны от импорта отдельных товаров, определяется по формуле

$$d_{Иj} = \frac{I_{ij}}{P_{ij}}$$

где I_{ij} – объем импорта i -го товара j -й страны; P_{ij} – объем потребления i -го товара j -й страны.

К оэффициент относительной экспортной специализации
коэффициент относительной экспортной специализации характеризует уровень международного разделения труда, определяется по формуле

$$K_{\text{оЭС}j} = \frac{d_{ij}}{d_i}$$

где d_{ij} – доля экспорта i -го товара j -й страны в общем объеме экспорта i -го товара; d_i – доля экспорта i -го товара в общем объеме экспорта

Коэффициент диверсификации

коэффициент диверсификации определяется по формуле

$$K_{\text{Див}} = \frac{\sum_{i=1}^n |d_{ij} - d_i|}{2}$$

- где n – объем товарной номенклатуры.

$K_{\text{Див}}$ определяется в интервале от 0 до 1.

Если он стремится к 1, значит страна специализируется на мировой рынок в производстве суженной номенклатуры товаров. При расширении производства экспортной продукции, реализуемой данной страной на мировом рынке, коэффициент диверсификации приближается к 0.

При $K_{\text{Див}} = 0$ структура экспорта страны абсолютно диверсифицирована, т.е. структура экспорта j -й страны совпадает с его мировой структурой



***ТЕМА 5. РЯДЫ
РАСПРЕДЕЛЕНИЯ В
ТАМОЖЕННОЙ
СТАТИСТИКЕ***

Ряды распределения в таможенной статистике

Признаки, разрабатываемые таможенной статистикой внешней торговли *варьируются* (отличаются друг от друга) у различных единиц совокупности в один и тот же период или момент времени.

Например, величина внешнеторгового оборота варьируется по подразделениям ФТС; величина экспорта (импорта) варьируется по направлениям экспорта (по разным странам-партнерам по внешней торговле), по видам товаров и т.п.

Причиной *вариации* являются разные условия существования разных единиц совокупности. Например, огромное число причин влияет на масштабы внешней торговли различных стран мира.

Ряды распределения в таможенной статистике

Первым этапом статистического изучения вариации является построение *ряда распределения (или вариационного ряда)* – упорядоченного распределения единиц совокупности по возрастающим (чаще) или по убывающим (реже) значениям признака и подсчет числа единиц с тем или иным значением признака.

Ряды распределения в таможенной статистике

Существует 3 вида ряда распределения:

- *ранжированный ряд;*
- *дискретный;*
- *интервальный ряд.*

Ранжированный ряд

Ранжированный ряд – это перечень отдельных единиц совокупности в порядке возрастания изучаемого признака;

Если численность единиц совокупности достаточно велика ранжированный ряд становится громоздким, и в таких случаях ряд распределения строится с помощью группировки единиц совокупности по значениям изучаемого признака

(если признак принимает небольшое число значений, то строится дискретный ряд, а в противном случае – интервальный ряд);

Дискретный ряд

дискретный ряд – это таблица, состоящая из двух столбцов (строк) – конкретных значений варьирующего признака X_i и числа единиц совокупности с данным значением признака f_i – частот;

число групп в дискретном ряду определяется числом реально существующих значений варьирующего признака;

Интервальный ряд

Интервальный ряд – это таблица, состоящая из двух столбцов (строк) – интервалов варьирующего признака X_i и числа единиц совокупности, попадающих в данный интервал (частот), или долей этого числа в общей численности совокупностей (частостей).

Этап I. Построение ранжированного ряда распределения

Построим ряд распределения внешнеторгового оборота (ВО) по таможенным постам России, для чего необходимо провести статистическое наблюдение, то есть собрать первичный статистический материал, который представляет собой величину ВО по всем таможенным постам, численность которых составляет 709 ед.

Этап I. Построение ранжированного ряда распределения

Ввиду огромного массива данных применение сплошного наблюдения экономически нецелесообразно, поэтому в таких случаях применяется *выборочный метод*, то есть из общего массива данных (генеральная совокупность) отбирается некоторая часть (выборочная совокупность, или выборка), которая и подвергается статистическому анализу.

При этом число единиц в выборке обозначают n , во всей генеральной совокупности – N .

Отношение n/N называется *относительный размер* или *частость выборки*.

Качество результатов выборочного метода зависит от *репрезентативности* выборки, т.е. от того, насколько она представительна в генеральной совокупности.

Для обеспечения репрезентативности выборки необходимо соблюдать принцип случайности отбора единиц.

Этап I. Построение ранжированного ряда распределения

В нашем примере про ВО примем частоту выборки $n/N = 0,05$ или 5%,

то есть в выборку включим $n = 0,05 * 709 = 35$ таможенных постов из 709.

Результаты выборочного наблюдения ВО по 35 таможенным постами за отчетный период представим в виде ранжированного по возрастанию величины ВО ряда распределения

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
млн.долл.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
МЛН.ДОЛЛ.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1		13			
2		14			
3		15			
4		16			
5		17			
6		18			
7		19			
8		20			
9		21			
10		22			
11		23			
12		24			

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
млн.долл.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1		13		25	
2		14		26	
3		15		27	
4		16		28	
5		17		29	
6		18		30	
7		19		31	
8		20		32	
9		21		33	
10		22		34	
11		23		35	
12		24		Итого	

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
МЛН.ДОЛЛ.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1	24,16	13		25	
2	27,06	14		26	
3	29,12	15		27	
4	31,17	16		28	
5	37,08	17		29	
6	39,11	18		30	
7	41,58	19		31	
8	44,84	20		32	
9	46,80	21		33	
10	48,37	22		34	
11	51,44	23		35	
12	52,56	24		Итого	

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
млн.долл.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1	24,16	13	54,12	25	
2	27,06	14	54,91	26	
3	29,12	15	55,74	27	
4	31,17	16	55,91	28	
5	37,08	17	56,07	29	
6	39,11	18	56,80	30	
7	41,58	19	56,93	31	
8	44,84	20	57,07	32	
9	46,80	21	58,39	33	
10	48,37	22	59,61	34	
11	51,44	23	59,95	35	
12	52,56	24	62,05	Итого	

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
млн.долл.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1	24,16	13	54,12	25	65,31
2	27,06	14	54,91	26	69,24
3	29,12	15	55,74	27	71,39
4	31,17	16	55,91	28	77,12
5	37,08	17	56,07	29	79,12
6	39,11	18	56,80	30	84,34
7	41,58	19	56,93	31	86,89
8	44,84	20	57,07	32	91,74
9	46,80	21	58,39	33	96,01
10	48,37	22	59,61	34	106,84
11	51,44	23	59,95	35	111,16
12	52,56	24	62,05	Итого	

Этап I. Построение ранжированного ряда распределения

Внешнеторговый оборот (ВО) по 35 таможенным постам,
млн.долл.

<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>	<i>№ поста</i>	<i>ВО</i>
1	24,16	13	54,12	25	65,31
2	27,06	14	54,91	26	69,24
3	29,12	15	55,74	27	71,39
4	31,17	16	55,91	28	77,12
5	37,08	17	56,07	29	79,12
6	39,11	18	56,80	30	84,34
7	41,58	19	56,93	31	86,89
8	44,84	20	57,07	32	91,74
9	46,80	21	58,39	33	96,01
10	48,37	22	59,61	34	106,84
11	51,44	23	59,95	35	111,16
12	52,56	24	62,05	Итого	2100,00

Этап I. Построение ранжированного ряда распределения

Учитывая, что на основе выборочного обследования нельзя точно оценить изучаемый параметр генеральной совокупности, необходимо найти пределы, в которых он находится.

Для этого необходимо определить изучаемый параметр по данным выборки

Этап I. Построение ранжированного ряда распределения

В нашем примере про ВО определим его средний размер в выборке, приняв за X величину ВО, а за N – численность выборки n :

$$\bar{X} = \frac{\sum X}{n} = 2100/35 = 60 \text{ (млн.долл.)}$$

Этап I. Построение ранжированного ряда распределения

Дисперсию определим по формуле:

$$\tilde{\sigma}^2 = \frac{\sum (X_i - \tilde{X})^2}{n} = \frac{(24,16 - 60)^2 + (27,06 - 60)^2 + \dots + (111,16 - 60)^2}{35} = 445,778 \text{ (млн.долл.}^2\text{)}$$

Затем необходимо определить *предельную ошибку выборки* по формуле

$$\Delta = t\mu \quad ,$$

где t – коэффициент доверия, зависящий от вероятности, с которой определяется предельная ошибка выборки;

μ – средняя ошибка выборки, определяемая для повторной выборки по формуле,

$$\mu = \sqrt{\frac{\tilde{\sigma}^2}{n}}$$

а для бесповторной – по формуле

$$\mu = \sqrt{\frac{\tilde{\sigma}^2}{n} \left(1 - \frac{n}{N}\right)}$$

Этап I. Построение ранжированного ряда распределения

В нашем примере про ВО выборка бесповторная, значит, применяя формулу, получим среднюю ошибку выборки при определении средней величины ВО в генеральной совокупности

$$\mu = \sqrt{\frac{445,778}{35} \left(1 - \frac{35}{709}\right)} = 3,48 \text{ (млн.долл.)}.$$

Значения вероятности P и коэффициента доверия t имеются в таблицах нормального закона распределения

Этап I. Построение ранжированного ряда распределения

Вероятность, которая принимается при расчете выборочной характеристики, называется *доверительной*.

Чаще всего принимают вероятность $P = 0,950$ ($t = 1,96$), которая означает, что только в 5 случаях из 100 ошибка может выйти за установленные границы.

Предельная ошибка выборки при определении средней величины ВО: $\Delta = 1,96 * 3,48 = 6,82$ (млн.долл.).

Этап I. Построение ранжированного ряда распределения

После расчета предельной ошибки находят *доверительный интервал* обобщающей характеристики генеральной совокупности для среднего значения, и для доли какого-либо признака:

$$\begin{array}{lll} \bar{X} = \tilde{X} \pm \Delta & \text{или} & (\tilde{X} - \Delta) \leq \bar{X} \leq (\tilde{X} + \Delta) \\ d = \tilde{d} \pm \Delta & \text{или} & (\tilde{d} - \Delta) \leq d \leq (\tilde{d} + \Delta) \end{array}$$

В нашем примере про ВО:

$\bar{X} = 60 \pm 6,82$ или $53,18 \leq \bar{X} \leq 66,82$ (млн.долл.), то есть средняя величина ВО в отчетном периоде по всем 709 таможенным постам с вероятностью 0,95 лежит в пределах от 53,18 млн.долл. до 66,18 млн.долл.

Этап 2. Построение интервального ряда распределения

Построим интервальный ряд распределения ВО по таможенным постам России, для чего необходимо выбрать оптимальное число групп (интервалов признака) и установить длину (размах) интервала.

Поскольку при анализе ряда распределения сравнивают частоты в разных интервалах, необходимо, чтобы длина интервалов была постоянной. Оптимальное число групп выбирается так, чтобы достаточной мере отразилось разнообразие значений признака в совокупности и в то же время закономерность распределении, его форма не искажалась случайными колебаниями частот. Если групп будет слишком мало, не проявится закономерность вариации; если групп будет чрезмерно много, случайные скачки частот исказят форму распределения.

Если приходится иметь дело с интервальным рядом распределения с неравными интервалами, то для сопоставимости нужно частоты или частости привести к единице интервала, полученное значение называется *плотностью* ρ , то есть $\rho = f/h$

Этап 2. Построение интервального ряда распределения

Чаще всего число групп в ряду распределения определяют по формуле Стержесса

$$k = 1 + 3,322 \lg N$$

$$k = 1 + 1,44 \ln N$$

где k – число групп (округляемое до ближайшего целого числа); N – численность совокупности

Этап 2. Построение интервального ряда распределения

Из формулы Стерджесса видно, что число групп – функция объема данных (N).

Зная число групп, рассчитывают длину (размах) интервала по формуле:

$$h = \frac{X_{\max} - X_{\min}}{k}$$

где X_{\max} и X_{\min} — максимальное и минимальное значения в совокупности.

Этап 2. Построение интервального ряда распределения

В нашем примере про ВО по формуле Стерждесса определим число групп:

$$k = 1 + 3,322/\lg 35 = 1 + 3,322 * 1,544 = 6,129 \approx 6.$$

Рассчитаем длину (размах) интервала по формуле:

$$h = (111,16 - 24,16)/6 = 87/6 = 14,5 \text{ (млн. долл.)}.$$

Этап 2. Построение интервального ряда распределения

Теперь построим интервальный ряд с 6 группами с интервалом 14,5 млн.долл. (см. первые 3 столбца табл.).

i	Группы постов по величине ВО X_i	Число постов f_i	Середина интервала X_i'	$X_i' f_i$	Накопл. частота f_i'	$ X_i' - \bar{X} f_i$	$(X_i' - \bar{X})^2 f_i$	$(X_i' - \bar{X})^3 f_i$	$(X_i' - \bar{X})^4 f_i$
1	24,16 – 38,66	5	31,41	157,05	5	147,071	4326,001	-127246,23	3742856,97
2	38,66 – 53,16	7	45,91	321,37	12	104,400	1557,051	-23222,31	346344,16
3	53,16 – 67,66	13	60,41	785,33	25	5,386	2,231	-0,92	0,38
4	67,66 – 82,16	4	74,91	299,64	29	56,343	793,629	11178,84	157461,90
5	82,16 – 96,66	4	89,41	357,64	33	114,343	3268,572	93434,47	2670891,13
6	96,66 – 111,16	2	103,91	207,82	35	86,171	3712,758	159966,81	6892284,32
	Итого	35		2128,85		513,714	13660,243	114110,66	13809838,86

Этап 2. Построение интервального ряда распределения

Существенную помощь в анализе ряда распределения и его свойств оказывает графическое изображение. **Интервальный ряд изображается столбиковой диаграммой, в которой основания столбиков, расположенные по оси абсцисс, – это интервалы значений варьирующего признака, а высоты столбиков – частоты, соответствующие масштабу по оси ординат.**

Этап 2. Построение интервального ряда распределения

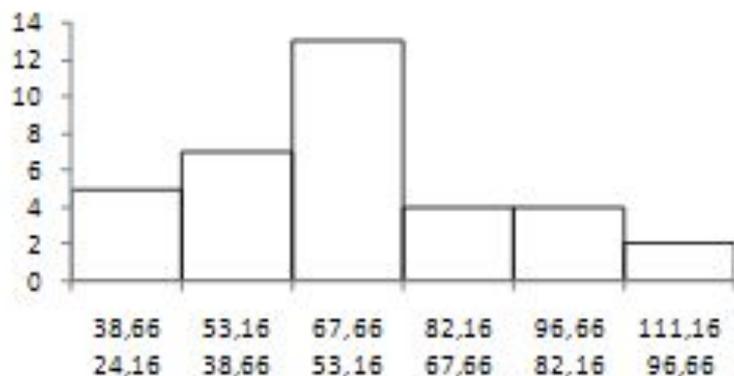


Рис. 8. Гистограмма распределения

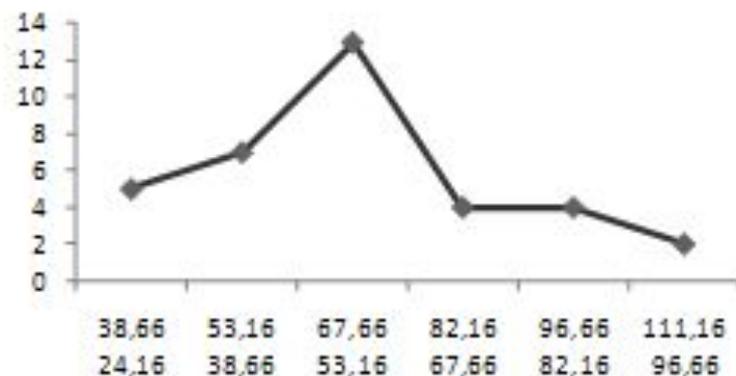


Рис. 9. Полигон распределения

Этап 2. Построение интервального ряда распределения

Вывод: чаще встречаются значения средних интервалов признака, реже – крайние (малые и большие) значения признака. Форма этого распределения близка к нормальному закону распределения, которое образуется, если на варьирующую переменную влияет большое число факторов, ни один из которых не имеет преобладающего значения.

Если имеется дискретный ряд распределения или используются середины интервалов (как в нашем примере про ВО – в таблице в 4-м столбце рассчитаны середины интервалов как полусумма значений начала и конца интервала), то графическое изображение такого ряда называется *полигоном* (см. рис. 9), которое получается соединением прямыми точек с координатами X_i и f_i .

Этап 3. Расчет структурных характеристик ряда

распределения

При изучении вариации применяются такие характеристики ряда распределения, которые описывают количественно его структуру, строение. Такова, например, *медиана* – величина варьирующего признака, делящая совокупность на две равные части – со значением признака меньше медианы и со значением признака больше медианы.

В нашем примере про ВО медиана – это 18-й таможенный пост из 35 с величиной ВО 56,8 млн.долл.

Из этого примера видно принципиальное различие между медианой и средней величиной: медиана не зависит от значений на краях ранжированного ряда. Даже если бы ВО 35-го таможенного поста был в 10 раз больше, величина медианы не изменилась бы.

Этап 3. Расчет структурных характеристик ряда распределения

Медиану часто используют как более надежный показатель типичного значения признака, нежели средняя арифметическая, если ряд значений неоднороден, включает резкие отклонения от средней. В интервальном ряду распределения для нахождения медианы применяется формула:

$$Me = X_0 + h \frac{0,5 \sum f - f'_{Me-1}}{f_{Me}}$$

Me – медиана;

X_0 – нижняя граница интервала, в котором находится медиана;

h – величина (размах) интервала;

f'_{Me-1} – накопленная частота в интервале, предшествующем медианному;

f_{Me} – частота в медианном интервале.

Этап 3. Расчет структурных характеристик ряда

распределения

Аналогично медиане вычисляются значения признака, делящие совокупность на 4 равные по численности части – *квартили*, которые обозначаются заглавной латинской буквой Q с подписным значком номера квартиля. Ясно, что Q_2 совпадает с Me . Для первого и третьего квартилей приводим формулы и расчет по данным табл. (в млн долл)

$$Q_1 = X_0 + h \frac{0,25 \sum f - f'_{Q_1-1}}{f_{Q_1}} = 38,66 + 14,5 \frac{0,25 * 35 - 5}{7} = 43,43$$

$$Q_3 = X_0 + h \frac{0,75 \sum f - f'_{Q_3-1}}{f_{Q_3}} = 67,66 + 14,5 \frac{0,75 * 35 - 25}{4} = 72,19$$

Этап 3. Расчет структурных характеристик ряда

распределения

Так как $Q_2 = Me = 59,30$ млн.долл., видно, что различие между первым квартилем и медианой ($-15,87$) больше, чем между медианой и третьим квартилем ($12,89$). Этот факт свидетельствует о наличии некоторой несимметричности в средней области распределения, что заметно и на рис. 8.

Значения признака, делящие ряд на 5 равных частей, называются *квинтилями*, на 10 частей – *децилями*, на 100 частей – *перцентилями*.

Эти характеристики применяются при необходимости подробного изучения структуры ряда распределения

Этап 3. Расчет структурных характеристик ряда распределения

Важное значение имеет такая величина признака, которая встречается в изучаемом ряду распределения чаще всего. Такую величину принято называть *модой*.

В дискретном ряду мода определяется без вычисления как значение признака с наибольшей частотой. Обычно встречаются ряды с одним модальным значением признака.

Если в ряду распределения встречаются 2 или несколько равных (и даже несколько различных, но больших чем соседние) значений признака, то он считается соответственно бимодальным или мультимодальным.

Это свидетельствует о неоднородности совокупности, возможно, представляющей собой агрегат нескольких совокупностей с разными модами.

Этап 3. Расчет структурных характеристик ряда распределения

В интервальном ряду распределения интервал с наибольшей частотой является модальным. Внутри этого интервала находят условное значение признака, вблизи которого *плотность распределения* (число единиц совокупности, приходящихся на единицу измерения варьирующего признака) достигает максимума.

Это условное значение и считается *точечной модой*.

Этап 3. Расчет структурных характеристик ряда

распределения

Точечная мода располагается ближе к той из границ интервала, за которой частота в соседнем интервале больше частоты в интервале за другой границей модального интервала. Отсюда получаем обычно применяемую формулу

$$M_o = X_0 + h \frac{f_{M_o} - f_{M_{o-1}}}{(f_{M_o} - f_{M_{o-1}}) + (f_{M_o} - f_{M_{o+1}})}$$

где M_o — мода;

X_0 — нижнее значение модального интервала;

f_{M_o} — частота в модальном интервале;

$f_{M_{o-1}}$ — частота в предыдущем интервале;

$f_{M_{o+1}}$ — частота в следующем интервале за модальным;

h — величина интервала

Этап 3. Расчет структурных характеристик ряда распределения

По данным табл. рассчитаем точечную моду по формуле

$$M_o = 53,16 + 14,5 \frac{13 - 7}{(13 - 7) + (13 - 4)} = 58,96$$

В интервальном ряду распределения ВО по таможенным постам средняя арифметическая рассчитывается как взвешенная по частоте середина интервалов X

$$\bar{X} = \frac{\sum X'f}{\sum f} = 2128,85/35 = 60,82 \text{ (млн. долл.)}$$

Различие между средней арифметической величиной (60,82), медианой (59,30) и модой (58,96) в нашем примере невелико. Чем ближе распределение по форме к нормальному закону, тем ближе значения медианы, моды и средней величины между собой.

Этап 4. Расчет показателей размера и интенсивности вариации

Простейшим показателем является *размах вариации* – абсолютная разность между максимальным и минимальным значениями признака из имеющихся в изучаемой совокупности значений .

$$H = X_{\max} - X_{\min}$$

Поскольку величина размаха характеризует лишь максимальное различие значений признака, она не может измерять закономерную силу его вариации во всей совокупности. Предназначенный для данной цели показатель должен учитывать и обобщать все различия значений признака в совокупности без исключения. Число таких различий равно числу сочетаний по два из всех единиц совокупности

Этап 4. Расчет показателей размера и интенсивности

вариации

Проще использовать среднюю из отклонений отдельных значений признака от среднего арифметического значения признака, а таковых в нашем примере про ВО всего 35. Но среднее отклонение значений признака от средней арифметической величины согласно первому свойству последней равно нулю. Поэтому **показателем силы вариации** выступает не арифметическая средняя отклонений, а **средний модуль отклонений, или среднее линейное отклонение**

$$L = \frac{\sum |X - \bar{X}|}{N}$$

Этап 4. Расчет показателей размера и интенсивности вариации

В нашем примере про ВО по данным табл. среднее линейное отклонение вычисляется как взвешенное по частоте отклонение по модулю средин интервалов от средней арифметической величины (расчет числителя произведен в 7-м столбце табл.), т.е. по формуле

$$L = \frac{\sum |X' - \tilde{X}| f}{\sum f} = 513,714 / 35 = 14,678$$

Это означает, что в среднем величина ВО в изучаемой совокупности таможенных постов отклонялась от средней величины ВО в РФ на 14,678 млн.долл.

Этап 4. Расчет показателей размера и интенсивности вариации

Простота расчета и интерпретации составляют положительные стороны показателя Λ , однако математические свойства модулей «плохие»: их нельзя поставить в соответствие с каким-либо вероятностным законом, в том числе и с нормальным распределением, параметром которого является не средний модуль отклонений, а *среднее квадратическое отклонение*, обозначаемое малой греческой буквой сигма (σ) или s и вычисляемое по формуле для ранжированного ряда и для интервального ряда

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

$$\sigma = \sqrt{\frac{\sum (X' - \bar{X})^2 f}{\sum f}}$$

Этап 4. Расчет показателей размера и интенсивности вариации

В нашем примере про ВО по данным табл. среднее квадратическое отклонение величины ВО по формуле составило (расчет числителя произведен в 8-м столбце табл)

$$\sigma = \sqrt{\frac{13660,243}{35}} = \sqrt{390,293} = 19,756$$

Этап 4. Расчет показателей размера и интенсивности вариации

Среднее квадратическое отклонение по величине в реальных совокупностях всегда больше среднего модуля отклонений. Разница между ними тем больше, чем больше в изучаемой совокупности резких, выделяющихся отклонений, что служит индикатором «засоренности» совокупности неоднородными с основной массой элементами. Для нормального закона распределения отношение $\sigma/L \approx 1,25$.

В нашем примере про ВО:

$$\sigma/L \approx 19,756 / 14,678 = 1,35 > 1,25$$

т.е. в изучаемой совокупности наблюдаются некоторое число таможенных постов с отличающимися от основной массы величинами ВО.

Этап 4. Расчет показателей размера и интенсивности вариации

Квадрат среднего квадратического отклонения представляет собой *дисперсию* отклонений, на использовании которой основаны практически все методы математической статистики, ее формула для несгруппированных данных (простая дисперсия) и для сгруппированных (взвешенная дисперсия):

$$\sigma^2 = \frac{\sum (X - \bar{X})^2}{n} = \overline{X^2} - \bar{X}^2$$

$$\sigma^2 = \frac{\sum (X' - \bar{X})^2 f}{\sum f} = \overline{X^2} - \bar{X}^2$$

Этап 4. Расчет показателей размера и интенсивности вариации

Еще одним показателем силы вариации, характеризующим ее не по всей совокупности, а лишь в ее центральной части, служит *среднее квартильное расстояние (отклонение)*, т.е. средняя величина разности между квартилями, определяемая по формуле

$$q = \frac{(Q_3 - Q_2) + (Q_2 - Q_1)}{2} = \frac{Q_3 - Q_1}{2}$$

В нашем примере про ВО по формуле

$$q = \frac{72,19 - 43,43}{2} = 14,38$$

Этап 4. Расчет показателей размера и интенсивности вариации

Сила вариации в центральной части совокупности, как правило, меньше, чем в целом по всей совокупности. Соотношение между средним линейным отклонением и средним квартильным расстоянием служит для изучения структуры вариации: большое значение такого соотношения свидетельствует о наличии слабоварьирующего «ядра» и сильно рассеянного вокруг него окружения в изучаемой совокупности.

Для нашего примера про ВО соотношение $L/q = 1,021$, что говорит о совсем незначительном различии силы вариации в центральной части совокупности и на ее периферии.

Этап 4. Расчет показателей размера и интенсивности вариации

Для оценки интенсивности вариации и для сравнения ее в разных совокупностях и тем более для разных признаков необходимы *относительные показатели вариации*, которые вычисляются как отношение абсолютных показателей силы вариации, рассмотренных ранее, к средней арифметической величине признака, то есть показатели:

относительный размах вариации:

$$\rho = \frac{H}{\bar{X}}$$

линейный коэффициент вариации:

$$\lambda = \frac{L}{\bar{X}}$$

квадратический коэффициент вариации:

$$v = \frac{\sigma}{\bar{X}}$$

относительное квартильное расстояние:

$$d = \frac{q}{\bar{X}}$$

В нашем примере про ВО эти показатели составляют:

$\rho = 87/60,82 = 1,43$, или 143%; $\lambda = 14,678/60,82 = 0,241$, или 24,1%;

$v = 19,756/60,82 = 0,32$, или 32%; $d = 14,38/60,82 = 0,236$, или 23,6%.

Этап 4. Расчет показателей размера и интенсивности вариации

Оценка степени интенсивности вариации возможна только для каждого отдельного признака и совокупности определенного состава, она состоит в сравнении наблюдаемой вариации с некоторой обычной ее интенсивностью, принимаемой за норматив.

Так, для совокупности таможенных постов вариация величины ВО может быть определена как слабая, если $< 25\%$, умеренная при $25\% < < 50\%$ и сильная при $> 50\%$.

Этап 5. Расчет моментов распределения и показателей его формы

Для дальнейшего изучения характера вариации используются средние значения разных степеней отклонений отдельных величин признака от его средней арифметической величины.

Эти показатели называются *центральные моменты распределения* порядка, соответствующего степени, в которую возводятся отклонения или просто моментов (нецентральные моменты в таможенной статистике практически не используются).

Этап 5. Расчет моментов распределения и показателей его формы

Центральные моменты

Порядок момента	Формула	
	по несгруппированным данным	по сгруппированным данным
Первый μ_1		
Второй μ_2		
Третий μ_3		
Четвертый μ_4		

Этап 5. Расчет моментов распределения и показателей его формы

Центральные моменты

Порядок момента	Формула	
	по несгруппированным данным	по сгруппированным данным
Первый μ_1	$\frac{\sum(x - \bar{x})}{N} = 0$	$\frac{\sum(x' - \bar{x})f}{\sum f} = 0$
Второй μ_2		
Третий μ_3		
Четвертый μ_4		

Этап 5. Расчет моментов распределения и показателей его формы

Центральные моменты

Порядок момента	Формула	
	по несгруппированным данным	по сгруппированным данным
Первый μ_1	$\frac{\sum(X - \bar{X})}{N} = 0$	$\frac{\sum(X' - \bar{X})f}{\sum f} = 0$
Второй μ_2	$\frac{\sum(X - \bar{X})^2}{N} = \bar{X}^2 - \bar{X}^2 = \sigma^2$	$\frac{\sum(X' - \bar{X})^2 f}{\sum f} = \bar{X}^2 - \bar{X}^2 = \sigma^2$
Третий μ_3		
Четвертый μ_4		

Этап 5. Расчет моментов распределения и показателей его формы

Центральные моменты

Порядок момента	Формула	
	по несгруппированным данным	по сгруппированным данным
Первый μ_1	$\frac{\sum(X - \bar{X})}{N} = 0$	$\frac{\sum(X' - \bar{X})f}{\sum f} = 0$
Второй μ_2	$\frac{\sum(X - \bar{X})^2}{N} = \bar{X}^2 - \bar{X}^2 = \sigma^2$	$\frac{\sum(X' - \bar{X})^2 f}{\sum f} = \bar{X}^2 - \bar{X}^2 = \sigma^2$
Третий μ_3	$\frac{\sum(X - \bar{X})^3}{N}$	$\frac{\sum(X' - \bar{X})^3 f}{\sum f}$
Четвертый μ_4		

Этап 5. Расчет моментов распределения и показателей его формы

Центральные моменты

Порядок момента	Формула	
	по несгруппированным данным	по сгруппированным данным
Первый μ_1	$\frac{\sum(X - \bar{X})}{N} = 0$	$\frac{\sum(X' - \bar{X})f}{\sum f} = 0$
Второй μ_2	$\frac{\sum(X - \bar{X})^2}{N} = \bar{X}^2 - \bar{X}^2 = \sigma^2$	$\frac{\sum(X' - \bar{X})^2 f}{\sum f} = \bar{X}^2 - \bar{X}^2 = \sigma^2$
Третий μ_3	$\frac{\sum(X - \bar{X})^3}{N}$	$\frac{\sum(X' - \bar{X})^3 f}{\sum f}$
Четвертый μ_4	$\frac{\sum(X - \bar{X})^4}{N}$	$\frac{\sum(X' - \bar{X})^4 f}{\sum f}$

Этап 5. Расчет моментов распределения и показателей его формы

Величина третьего момента μ^3 зависит, как и его знак, от преобладания положительных кубов отклонений над отрицательными кубами либо наоборот. При нормальном и любом другом строго симметричном распределении сумма положительных кубов строго равна сумме отрицательных кубов, поэтому на основе третьего момента строится показатель, характеризующий степень асимметричности распределения – *коэффициент асимметрии*

$$As = \frac{\mu_3}{\sigma^3}$$

Этап 5. Расчет моментов распределения и показателей его формы

В нашем примере про ВО показатель асимметрии составил (расчет числителя произведен в 9-м столбце табл.)

$$As = \frac{114110,66}{19,756^3 * 35} = 0.423 > 0$$

т.е. симметрия значительна

Этап 5. Расчет моментов распределения и показателей его формы

Английский статистик К.Пирсон на основе разности между средней арифметической величиной и модой предложил другой показатель асимметрии

$$As_{\Pi} = \frac{\bar{X} - Mo}{\sigma}$$

В нашем примере по данным табл. показатель асимметрии составил

$$As = \frac{60,82 - 58,96}{19,756} = 0.09$$

Этап 5. Расчет моментов распределения и показателей его формы

Показатель асимметрии Пирсона зависит от степени асимметричности в средней части ряда распределения, а показатель асимметрии – от крайних значений признака.

Этап 5. Расчет моментов распределения и показателей его формы

Таким образом, в нашем примере про ВО в средней части распределения наблюдается меньшая асимметрия, чем по краям, что видно и по графику (рис. 9). Распределения с сильной правосторонней и левосторонней асимметрией показаны на рис. 10.

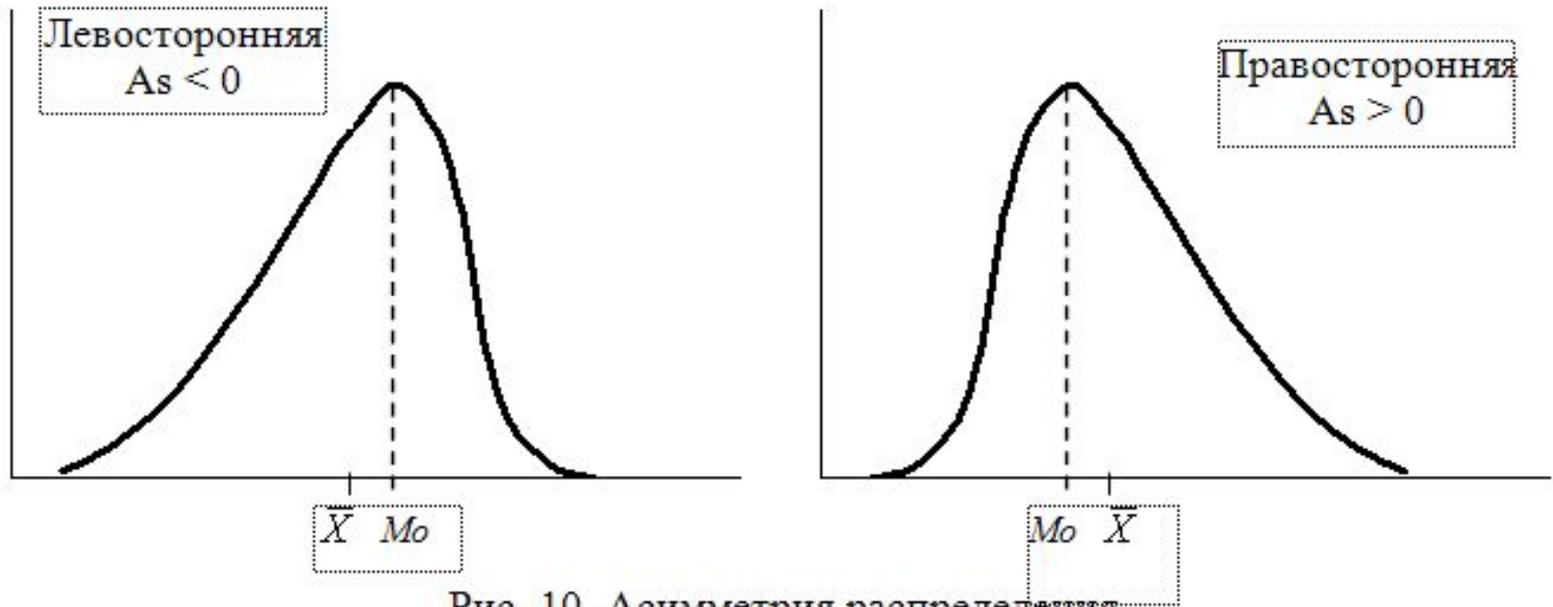


Рис. 10. Асимметрия распределения

Этап 5. Расчет моментов распределения и показателей его формы

С помощью момента четвертого порядка характеризуется еще более сложное свойство рядов распределения – *эксцесс* (от англ. «излишество»). Показатель эксцесса рассчитывается по формуле

$$Ex = \frac{\mu_4}{\sigma^4} - 3$$

Этап 5. Расчет моментов распределения и показателей его формы

Чаще всего эксцесс интерпретируется как «крутизна» распределения, что не совсем верно. График распределения может выглядеть сколь угодно крутым в зависимости от силы вариации признака: чем слабее вариация, тем круче кривая распределения при данном масштабе.

Этап 5. Расчет моментов распределения и показателей его формы

Чтобы показать, в чем состоит эксцесс распределения, и правильно его интерпретировать, нужно сравнить ряды с одинаковой силой вариации (одной и той же величиной σ) и разными показателями эксцесса. Чтобы не смешать эксцесс с асимметрией, все сравниваемые ряды должны быть симметричными. Такое сравнение изображено на рис. 11

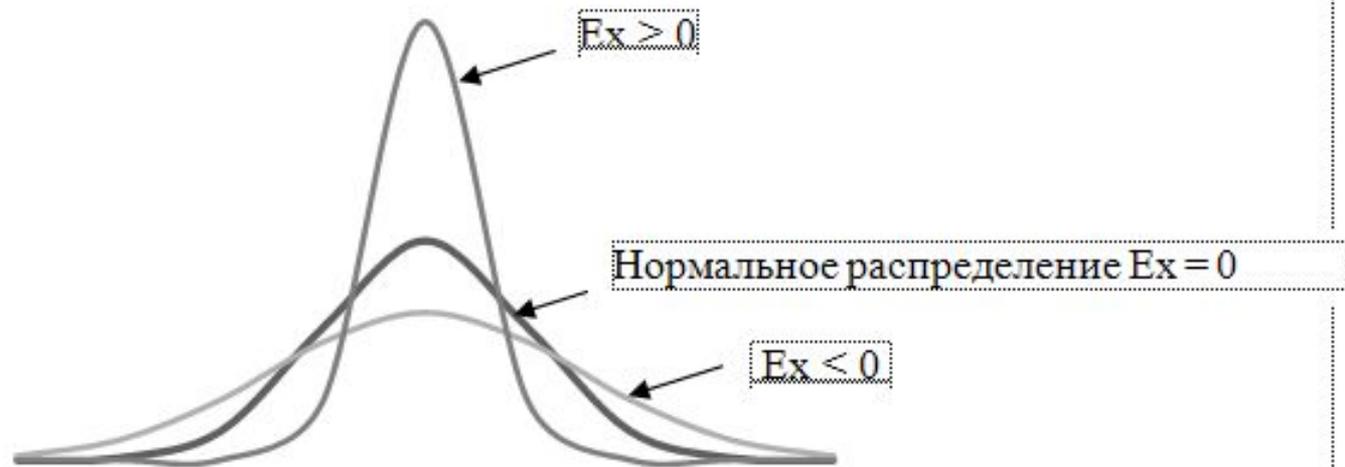


Рис. 11. Эксцесс распределения

Этап 5. Расчет моментов распределения и показателей его формы

Наличие положительного эксцесса означает наличие слабоварьирующего «ядра» и сильно рассеянного вокруг него окружения в изучаемой совокупности. Отрицательный эксцесс означает отсутствие такого «ядра».

В нашем примере эксцесс составил

$$E_x = \frac{13809838,86}{19,756^4 * 35} - 3 = -0,41$$

(расчет числителя произведен в 10-м столбце табл. 18): , т.е. величина ВО по таможенным постам варьирует сильнее, чем при нормальном распределении.

Этап 5. Расчет моментов распределения и показателей его формы

По значениям показателей асимметрии и эксцесса распределения можно судить о близости распределения к нормальному: показатели асимметрии и эксцесса не должны превышать своих двукратных средних квадратических отклонений, т.е.

$$|As| < 2\sigma_{As}$$

$$|Ex| < 2\sigma_{Ex}$$

Этап 5. Расчет моментов распределения и показателей его формы

Эти средние квадратические отклонения
вычисляются по формулам

$$\sigma_{As} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$$

$$\sigma_{Ex} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$$

Этап 5. Расчет моментов распределения и показателей его формы

В нашем примере по формулам (57) и (58):

$$\sigma_{As} = \sqrt{\frac{6 * 35 * (35 - 1)}{(35 - 2)(35 + 1)(35 + 3)}} = 0,40 \quad \sigma_{Ex} = \sqrt{\frac{24 * 35 * (35 - 1)^2}{(35 - 3)(35 - 2)(35 + 3)(35 + 5)}} = 0,78$$

Так как показатели асимметрии и эксцесса не превышают своих двукратных средних квадратических отклонений ($As = |0,423| < 0,4 * 2$; $Ex = |-0,41| < 0,78 * 2$), можно говорить о сходстве анализируемого распределения с нормальным.

Этап 6. Проверка соответствия ряда распределения теоретическому

Под теоретической кривой распределения понимается графическое изображение ряда в виде непрерывной линии изменения частот в вариационном ряду, функционально связанного с изменением вариантов, другими словами, теоретическое распределение может быть выражено аналитически – формулой, которая связывает частоты и соответствующие значения признака. Такие алгебраические формулы носят название *законов распределения*.

Большое познавательное значение имеет сопоставление фактических кривых распределения с теоретическими.

Этап 6. Проверка соответствия ряда распределения теоретическому

часто пользуются типом распределения, которое называется *нормальным*. Формула функции плотности нормального распределения имеет следующий вид:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\bar{X})^2}{2\sigma^2}} \quad \varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

где X – значение изучаемого признака;

– \bar{X} – средняя арифметическая ряда;

σ – среднее квадратическое отклонение;

$t = \frac{X - \bar{X}}{\sigma}$ – нормированное отклонение;

$\pi = 3,1415$ – постоянное число (отношение длины окружности к ее диаметру);

$e = 2,7182$ – основание натурального логарифма.

Этап 6. Проверка соответствия ряда распределения теоретическому

- Следовательно, кривая нормального распределения может быть построена по двум параметрам – средней арифметической и среднему квадратическому отклонению. Поэтому важно выяснить, как эти параметры влияют на вид нормальной кривой.
- Если \bar{X} не меняется, а изменяется только σ , то чем меньше σ , тем более вытянута вверх кривая и наоборот, чем больше σ , тем более плоской и растянутой вдоль оси абсцисс становится кривая нормального распределения

Этап 6. Проверка соответствия ряда распределения теоретическому

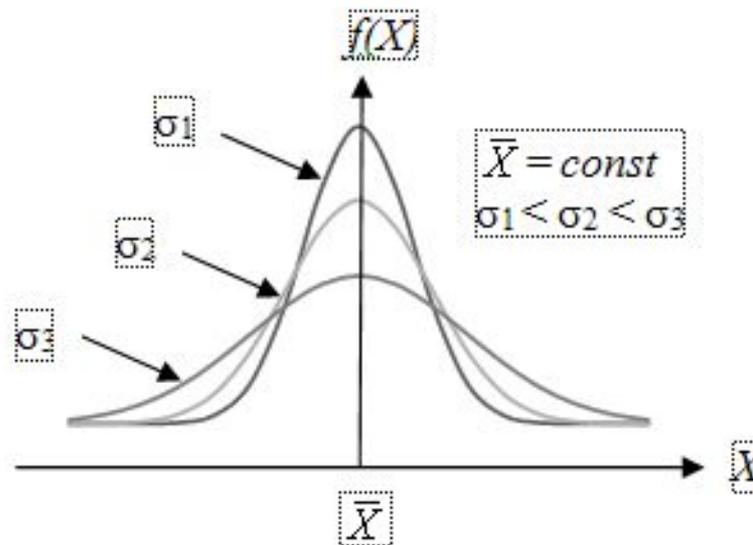


Рис. 12. Влияние величины σ на кривую нормального распределения

Этап 6. Проверка соответствия ряда распределения теоретическому

Если σ остается неизменной, а \bar{X} изменяется, то кривые нормального распределения имеют одинаковую форму, но отличаются друг от друга положением максимальной ординаты (вершины)

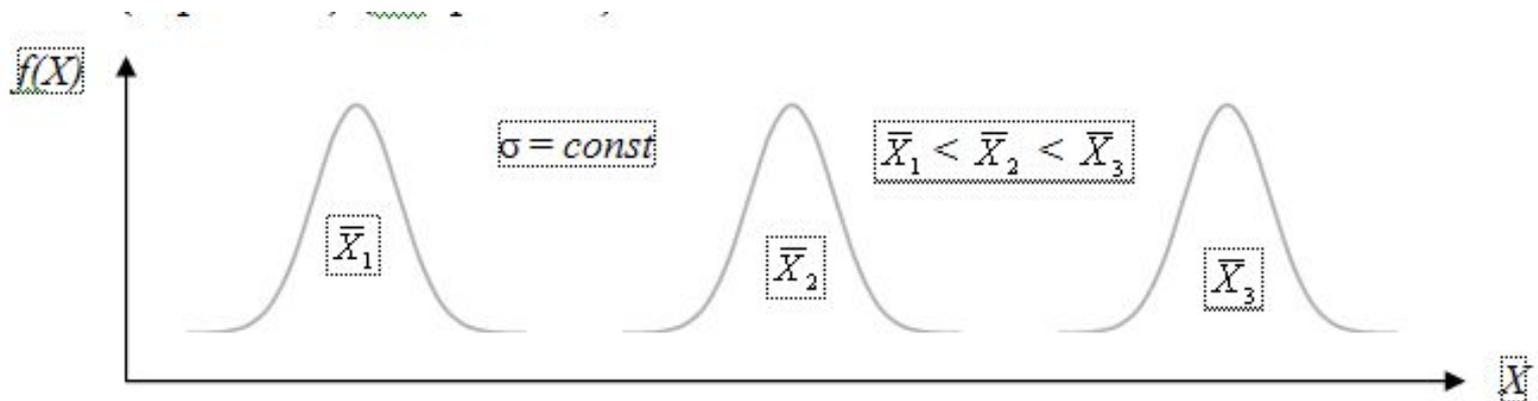


Рис. 13. Влияние величины \bar{X} на кривую нормального распределения

Этап 6. Проверка соответствия ряда распределения теоретическому

особенности кривой нормального
распределения

- кривая симметрична и имеет максимум в
точке, соответствующей значению $\bar{X} = Me =$

Этап 6. Проверка соответствия ряда распределения теоретическому

особенности кривой нормального распределения

- кривая симметрична и имеет максимум в точке, соответствующей значению $\bar{X} = Me = Mo$;
- кривая асимптотически приближается к оси абсцисс, продолжаясь в обе стороны до бесконечности (чем больше \bar{X} отдельные значения X отклоняются от \bar{X} , тем реже они

Этап 6. Проверка соответствия ряда распределения теоретическому

особенности кривой нормального распределения

- кривая симметрична и имеет максимум в точке, соответствующей значению $\bar{X} = Me = Mo$;
- кривая асимптотически приближается к оси абсцисс, продолжаясь в обе стороны до бесконечности (чем больше \bar{X} отдельные значения X отклоняются от \bar{X} , тем реже они встречаются);
- кривая имеет две точки перегиба на

Этап 6. Проверка соответствия ряда распределения теоретическому

особенности кривой нормального распределения

- кривая симметрична и имеет максимум в точке, соответствующей значению $\bar{X} = Me = Mo$;
- кривая асимптотически приближается к оси абсцисс, продолжаясь в обе стороны до бесконечности (чем больше \bar{X} отдельные значения X отклоняются от \bar{X} , тем реже они встречаются);
- кривая имеет две точки перегиба на расстоянии $\pm \sigma$ от \bar{X} ;
- коэффициенты асимметрии и эксцесса равны нулю

Этап 6. Проверка соответствия ряда распределения теоретическому

Проверка гипотезы о соответствии теоретическому распределению предполагает расчет теоретических частот этого распределения.

Для нормального распределения порядок расчета этих частот следующий:

по эмпирическим данным рассчитывают среднюю арифметическую ряда \bar{X} и среднее квадратическое отклонение σ ;

находят нормированное (выраженное в σ) отклонение каждого эмпирического значения от средней арифметической:

$$t = \frac{X - \bar{X}}{\sigma}$$

Этап 6. Проверка соответствия ряда распределения теоретическому

- по формуле или с помощью таблиц интеграла вероятностей Лапласа находят значение $\phi(t)$
- вычисляют теоретические частоты m по формуле:

$$m_i = Nh_i\varphi(t)$$

Этап 6. Проверка соответствия ряда распределения теоретическому

- Критерии согласия, опираясь на установленный закон распределения, дают возможность установить, когда расхождения между теоретическими и эмпирическими частотами следует признать несущественными (случайными), а когда – существенными (неслучайными). Таким образом, критерии согласия позволяют отвергнуть или подтвердить правильность выдвинутой гипотезы о характере распределения в эмпирическом ряду и дать ответ, можно ли принять для данного эмпирического распределения модель, выраженную некоторым теоретическим законом распределения.

Существует ряд критериев согласия, но чаще всего применяют критерии Пирсона χ^2 , Колмогорова и Романовского.

Этап 6. Проверка соответствия ряда распределения теоретическому

Критерий согласия Пирсона χ^2 (хи-квадрат) – один из основных критериев согласия, рассчитываемый по формуле:

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - m_i)^2}{m_i}$$

k – число интервалов;

f_i – эмпирическая частота i -го интервала;

m_i – теоретическая частота.

Для распределения χ^2 составлены таблицы, где указано критическое значение критерия согласия χ^2 для выбранного уровня значимости α и данного числа степеней свободы ν

Этап 6. Проверка соответствия ряда распределения теоретическому

Уровень значимости α – это вероятность ошибочного отклонения выдвинутой гипотезы, т.е. вероятность (P) того, что будет отвергнута правильная гипотеза.

В статистических исследованиях в зависимости от важности и ответственности решаемых задач пользуются следующими тремя уровнями значимости:

$\alpha = 0,10$, тогда $P = 0,90$;

$\alpha = 0,05$, тогда $P = 0,95$;

$\alpha = 0,01$, тогда $P = 0,99$.

Практически приемлемая вероятность в экономических исследованиях, означающая, что в 5 случаях из 100 может быть отвергнута правильная гипотеза

Этап 6. Проверка соответствия ряда распределения теоретическому

Число степеней свободы ν определяется по формуле:

$$\nu = k - z - 1,$$

где k – число интервалов;

z – число параметров, задающих теоретический закон распределения.

Для нормального распределения $z = 2$, так как нормальное распределение зависит от двух параметров – средней арифметической (μ) и среднего квадратического отклонения (σ).

Этап 6. Проверка соответствия ряда распределения теоретическому

Для оценки существенности расхождений расчетное значение χ^2 сравнивают с табличным $\chi^2_{\text{табл}}$.

Расчетное значения критерия должно быть меньше табличного, т.е. $\chi^2 < \chi^2_{\text{табл}}$, в противном случае расхождения между теоретическим и эмпирическим распределением не случайны, а теоретическое распределение не может служить моделью для изучаемого эмпирического распределения.

Этап 6. Проверка соответствия ряда распределения теоретическому

Использование критерия χ^2 рекомендуется для достаточно больших совокупностей ($N > 50$), при этом частота каждой группы не должна быть менее 5, в противном случае повышается вероятность получения ошибочных выводов.

Этап 6. Проверка соответствия ряда распределения теоретическому

Критерий Романовского K_P основан на использовании критерия Пирсона χ^2 , т. е. уже найденных значений χ^2 и числа степеней свободы ν , рассчитывается по формуле

$$K_P = \frac{|\chi^2 - \nu|}{\sqrt{2\nu}}$$

Этап 6. Проверка соответствия ряда распределения теоретическому

Он используется в том случае, когда отсутствует таблица значений χ^2 .

Если $K_p < 3$, то расхождения между теоретическим и эмпирическим распределением случайны, если $K_p > 3$, то не случайны, и теоретическое распределение не может служить моделью для изучаемого эмпирического распределения

Этап 6. Проверка соответствия ряда распределения теоретическому

Критерий Колмогорова λ основан на определении максимального расхождения между накопленными частотами эмпирического и теоретического распределений (D), рассчитывается по формуле

$$\lambda = D / \sqrt{N}$$

Этап 6. Проверка соответствия ряда распределения теоретическому

Рассчитав значение λ , по таблице $P(\lambda)$ определяют вероятность, с которой можно утверждать, что отклонения эмпирических частот от теоретических случайны.

Вероятность $P(\lambda)$ может изменяться от 0 до 1. При $P(\lambda) = 1$ (т.е. при $\lambda < 0,3$) происходит полное совпадение частот, при $P(\lambda) = 0$ – полное расхождение.

Вывод

Итак, подтвердив правильность выдвинутой гипотезы с помощью известных критериев согласия, можно использовать результаты распределения для практической деятельности.

- Во-первых, соответствие нормальному закону позволяет прогнозировать, какое число таможенных постов (или их доля) попадет в тот или иной интервал значений величины ВО.
- Во-вторых, нормальное распределение возникает при действии на вариацию изучаемого показателя множества независимых факторов.
- Из чего следует, что нельзя существенно снизить вариацию величины ВО, воздействуя только на один-два управляемых фактора, скажем число работников таможенного поста или степень технической оснащенности.