
НОВЫЕ ПОДХОДЫ К СТАТИСТИЧЕСКОМУ РАСЧЕТУ И АНАЛИЗУ МАКРОЭКОНОМИЧЕСКИХ ИНДИКАТОРОВ

Дарманян Анатолий Петрович,

доктор технических наук, профессор кафедры математики и информатики Волгоградского филиала Финансового университета при Правительстве РФ; adarma@inbox.ru

На основе расчета статистических критериев эмпирических распределений макроэкономических индикаторов РФ и стран ЕС показана целесообразность использования других показателей описательной статистики вместо средних арифметических значений, а в качестве графической иллюстрации результатов статистических расчетов – диаграммы типа «ящик с усами».

Ключевые слова: макроэкономические индикаторы, описательная статистика, нормальное распределение, диаграмма «ящик с усами».

При проведении анализа социально-экономического состояния страны и отдельных ее регионов с помощью различных макроэкономических индикаторов возникают определенные трудности, связанные с тем, что эти индикаторы получены как усредненные значения по эмпирическим выборкам Росстата [5]. Традиционно макроэкономические индикаторы рассчитываются методами описательной статистики, когда в качестве центральной тенденции эмпирического распределения используется среднее арифметическое значение и реже доверительный интервал, что именуется в Росстате как «в среднем за период» [5]. Анализ социально-экономического состояния и развития как РФ в целом, так и отдельных ее регионов при таком подходе сводится, как правило, к сравнению средних арифметических значений макроэкономических показателей для одного временного периода по различным качественным критериям (субъекты РФ, регионы, отрасли производства, виды деятельности и т.д.), а для разных временных периодов – как процентное изменение выбранного макроэкономического индикатора («в % к соответствующему периоду предыдущего года» [5]).

Однако использование среднеарифметических значений правомерно только в том случае, когда анализируемое распределение данных имеет нормальное распределение (распределение Гаусса), при котором 68,3% всех значений находится в интервале: среднее значение плюс (минус) стан-

дартное отклонение [1, с. 188]. Только в этом случае такие показатели описательной статистики как среднее арифметическое значение эмпирического распределения вместе с доверительным интервалом могут быть использованы в качестве макроэкономического индикатора, а его процентное изменение как характеристика различия или тенденция изменения. И только в этом случае можно проводить сравнение макроэкономических показателей в различных регионах с целью выявления различий между ними.

В случае же отсутствия нормального распределения более правильно (и как будет показано далее, гораздо более информативно) использовать другие показатели описательной статистики, такие как медиана, минимальное и максимальное значение показателя, интерквартильный размах (IQR, разница значений верхнего 75-го нижнего 25-го квартилей) [4, с. 71].

Поэтому целью данной работы являлось доказательство неправомерности использования средних арифметических значений эмпирических распределений макроэкономических индикаторов и иллюстрация возможностей для экономического анализа, которые предоставляют другие показатели описательной статистики и их графическое представление в виде диаграммы «ящик с усами» («Box&WhiskerPlot») [1, с. 159; 4, с. 57].

В научном плане такой подход совсем не нов и уже много лет строго регламентирован для зарубежных медиков-исследователей, использующих медицинскую статистику для обработки эмпирических данных, распределения которых отличаются (как и в экономике) большим разбросом значений и малыми выборками [2, с. 36; 3, с. 31].

Для настоящего исследования были выбраны следующие макроэкономические индикаторы: среднедушевые доходы населения РФ, уровень занятости в РФ (доля занятых в экономически активном населении, %), уровень занятости в странах Евросоюза (%).

Для проверки соответствия этих распределений нормальному закону распределения рассчитывались следующие статистические критерии: критерий Колмогорова-Смирнова (K-Stest) и критерий Шапиро-Уилки (SW-Wtest) [3, с. 42]. В качестве эмпирических распределений для РФ были использованы данные Росстата [5], а для стран Евросоюза (ЕС) – данные Евростата [7]. Все расчеты выполнены с помощью пакета прикладных программ Statisticav.6.0 [6].

В табл. 1 приведены результаты расчетов статистических критериев для всех выбранных эмпирических распределений макроэкономических индикаторов с целью проверки соответствия их нормальному закону распределения (распределению Гаусса). На рис. 1 приведена гистограмма распределения среднедушевых доходов населения 83 регионов РФ, а на рис. 2 – распределение уровня занятости населения в 27 странах Евросоюза, на которых также показан график нормального распределения.

Проверка нормальности распределений макроэкономических индикаторов для уровня значимости $p=0,05$

Показатель	Критерий K-S	Критерий SW-W
Среднедушевые денежные доходы населения 83 регионов РФ, (руб. в месяц)	0,222	0,692
Уровень занятости в субъектах РФ в 2011 г., (%)	0,840	0,152
Уровень занятости в 27 странах ЕС в 2011 г., (%)	0,165	0,767

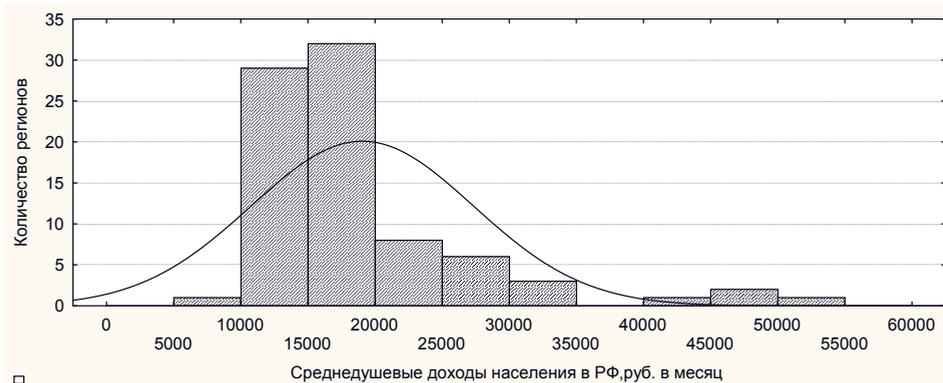


Рис. 1. Гистограмма распределения среднедушевых доходов в 83 субъектах РФ в 2011 г. (сплошная линия – нормальное распределение)

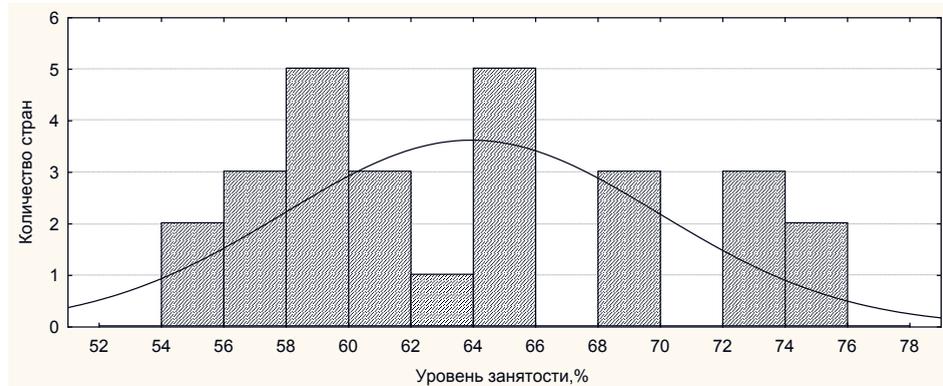


Рис. 2. Гистограмма распределения уровня занятости населения в 27 странах ЕС в 2011 г. (сплошная линия – нормальное распределение)

Как ясно видно из рис. 1 и рис. 2, распределение среднедушевых доходов населения во всех регионах РФ и уровень занятости в странах ЕС не являются нормальными распределениями. Этот же вывод подтверждают и расчеты критерия Колмогорова – Смирнова (K-S) и критерия Шапиро – Уилки (SW-W) (табл. 1).

Так как нормальное распределение для уровня значимости $p=0,05$ (95% уровень достоверности) имеет место тогда, когда $K-S \leq 0,0895$ и $SW-W \geq 0,947$ [3, с. 42], то из данных табл. 1 видно, что все анализируемые эмпирические

распределения не являются нормальными распределениями и к ним такие показатели описательной статистики, как среднее арифметическое значение, стандартное отклонение, доверительный интервал, строго говоря, не применимы.

Особенно это наглядно видно на гистограмме рис. 2. Действительно, о какой средней занятости населения в странах ЕС (в 2011 г. она составляла «в среднем» 64%) можно говорить при таком распределении?

При отсутствии нормального распределения данных необходимо рассчитывать, как отмечалось выше, другие показатели описательной статистики, а в качестве графической иллюстрации распределений целесообразно использовать диаграммы типа «ящик с усами» («Box&WhiskerPlot»), которые обеспечивают как диагностическую, так и описательную информацию об исследуемой совокупности эмпирических данных. Поэтому в табл. 2 приведены результаты расчетов различных показателей описательной статистики для всех выбранных эмпирических распределений.

Таблица 2

Показатели описательной статистики эмпирических распределений макроэкономических индикаторов

Индикатор	Среднее	Медиана	Нижний квартиль	Верхний квартиль	Размах, IQR
Денежные доходы, руб. в месяц	18931	16032	14353	20146	5793
Уровень занятости в РФ, %	61,4	62,0	59,4	64,1	4,7
Уровень занятости в ЕС, %	63,9	63,8	58,5	69,0	10,5

Рассчитанные статистические показатели в табл. 2 предоставляют исследователю дополнительную информацию для проведения анализа распределений эмпирических данных. Так, медиана соответствует центру распределения и лучше, чем среднее арифметическое значение, характеризует эмпирическое распределение при ассиметричном распределении и малом числе значений в выборке. Так как квартили делят выборку на четыре равные части, то интерквартильный размах (IQR), в котором располагается 50% всех данных, дает ясное представление о ширине распределения и с экономической точки зрения количественно характеризует социально-экономические отличия разных регионов.

Еще одно несомненное преимущество таких статистических показателей, как медиана и квартили, заключается в том, что они не подвержены влиянию выбросов и экстремальных значений, которые всегда имеют место в эмпирическом распределении любого макроэкономического показателя в такой огромной и разноплановой в социальном и экономическом смысле стране как Россия. То же самое справедливо и для всех 27 стран Евросоюза (рис. 2).

Для наглядности на рис. 3 и рис. 4 показаны диаграммы типа «ящик с усами» для двух эмпирических распределений. Как видно из данных в табл. 2 и на графике рис. 3, в 2011 г. в половине субъектов РФ (41 Федеральный округ, ФО) среднедушевые доходы населения были меньше 16032 руб. в месяц, а в половине (41 ФО) – больше этого размера (значение медианы, median). Кроме того, в половине округов (41 ФО) среднедушевые доходы находились в интервале значений от 14353 руб. до 20145 руб. в месяц (интерквартильный размах, IQR), в то время «в среднем за период» среднедушевые доходы населения «в среднем» для всех субъектов РФ составляли 18931 руб. в месяц (табл. 2).

На диаграмме рис.3 обращают на себя внимание большие «усы» (Non-OutlierRange), которые свидетельствуют о широком разбросе среднедушевых доходов населения в субъектах РФ (от 8829 до 25994 руб. в месяц). Также на рис. 3 видны выбросы (Outliers) и значительные экстремальные значения (Extremes) среднедушевых доходов в субъектах РФ.

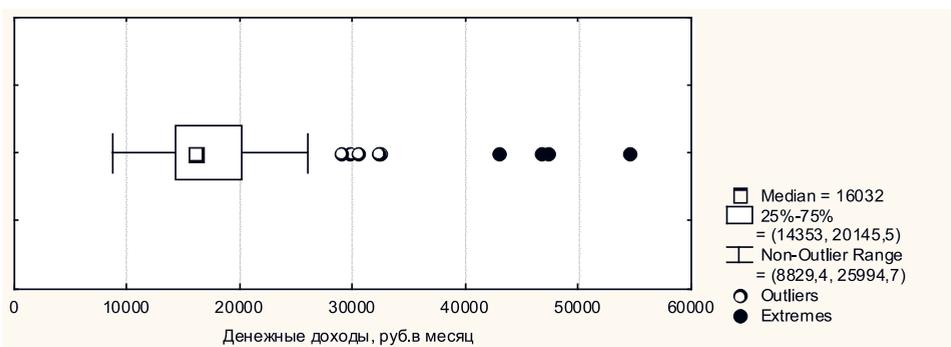


Рис. 3. Диаграмма «ящик с усами» распределения среднедушевых доходов в 83 субъектах РФ в 2011 г.

Эти значения свидетельствуют о том, что в РФ есть субъекты, в которых среднедушевые доходы населения находятся в пределах 30000 руб. в месяц, а есть и регионы с денежными доходами населения от 43000 до 55000 руб. в месяц. Это лишний раз подтверждает факт большого разброса эмпирических данных среднедушевого дохода в регионах РФ, что и оказывает сильное влияние на вычисление средних арифметических значений.

Интересные и полезные выводы можно сделать на основании анализа диаграммы на рис. 4, на которой показано распределение занятости населения в субъектах РФ. По значению медианы можно судить о том, что в половине субъектов РФ (41 ФО) занятость населения была меньше 62%, а в половине (41 ФО) – больше 62% (табл. 2). Кроме того, в половине округов занятость населения находилась в интервале значений от 59,4% до 64,1% (интерквартильный размах в табл. 2), в то время как «в среднем за период» уровень занятости для всех субъектов РФ составлял (табл. 2).

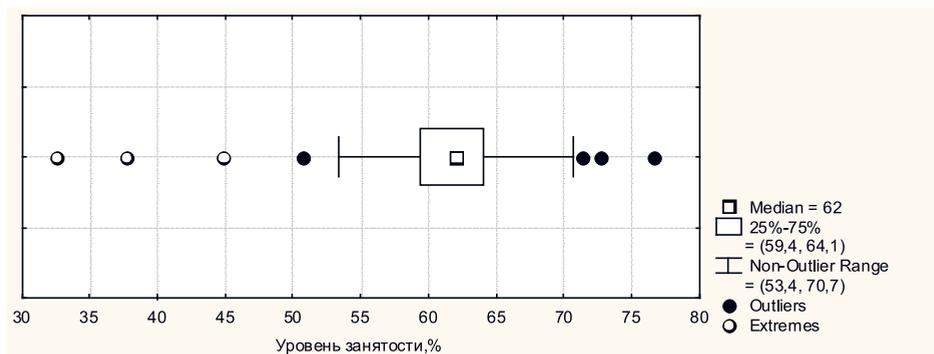


Рис. 4. Диаграмма «ящик с усами» распределения уровня занятости в 83 субъектах РФ в 2011 г.

На диаграмме рис. 4 обращают на себя внимание большие «усы» (Non-OutlierRange), которые свидетельствуют о широком разбросе уровня занятости населения в субъектах РФ (от 53,4% до 70,7%). Также на рис. 4 видны многочисленные выбросы (Outliers) и экстремальные значения (Extremes), которые свидетельствуют о большом разбросе уровня занятости населения (от 32% до 76%) в субъектах РФ и являются причиной некорректных выводов об этом макроэкономическом показателе для РФ в целом, как «в среднем».

На основе выполненного исследования можно сделать следующие выводы:

1. Так как отдельные эмпирические распределения макроэкономических показателей не отвечают закону нормального распределения, то для их характеристики неправомерно использовать средние арифметические значения.
2. Более информативно и правильно в таких случаях использовать такие показатели описательной статистики, как медиана и интерквартильный размах, а в качестве графической иллюстрации – диаграммы типа «ящик с усами».

Список источников

1. Берк, К. Анализ данных с помощью Microsoft Excel [текст] / К. Берк, П. Кейри / Пер с англ. – Издательский дом «Вильямс», 2005. – 560 с.
2. Ланг, Т.А. Как описывать статистику в медицине [текст] / Т.А. Ланг, М. Сесик // Аннотированное руководство для авторов, редакторов и рецензентов / Пер. с англ. под ред. В.П. Леонова. – М.: Практическая медицина, 2011. – 480 с.
3. Мамаев, А.Н. Основы медицинской статистики [текст] / А.Н. Мамаев. – М.: Практическая медицина, 2011. – 128 с.
4. Методы анализа распределений. Выборочное наблюдение: учеб. пособие [текст] / Н.В. Куприенко и др. – 3-е изд. – СПб.: Изд-во Политехн. ун-та, 2009. – 138 с.
5. Российский статистический ежегодник. 2011: Стат. сб. [текст] // Росстат. – М., 2011. – 795 с.

6. Халафян, А.А. STATISTICA 6. Статистический анализ данных, 3-е изд. учебник [текст] / А.А. Халафян. – М.: ООО «Бином-Пресс», 2007. – 512 с.
7. EU employment statistics [электронный ресурс]. – URL: http://ec.europa.eu/eurostat.ec.europa.eu/statistics_explained//index.

NEW APPROACHES TO STATISTICAL CALCULATION AND ANALYSES OF MACROECONOMICAL INDICATORS

Darmanyam Anatoliy Petrovich,

Dr. Sc. of Technical Sciences, Professor of the Chair of Mathematics and Computer Science of Volgograd filial-branch of Financial university under the Government of the Russian Federation; adarma@inbox.ru

On the basis of the calculation of statistical criteria of the empirical distributions of the macroeconomic indicators of the Russian Federation and the countries of the EU shows the expediency of the use of other indicators descriptive statistics instead of mean, and as a graphic illustration of the calculation results – chart «Box & Whisker Plot».

Keywords: economic indicators, descriptive statistics, normal distribution, chart «Box-and-Whisker Plot».

ПРАВИЛА ДЛЯ АВТОРОВ

Журнал «Современная экономика: проблемы и решения» принимает к публикации материалы, содержащие результаты оригинальных исследований, оформленных в виде полных статей (до 20 страниц) и кратких сообщений (до 5 страниц).

Опубликованные материалы, а также материалы, представленные для публикации в других журналах, к рассмотрению не принимаются.

Для публикации авторы предоставляют следующие материалы в редакцию журнала (по электронной почте: journal.MEPR@yandex.ru):

1. **Статью**, набранную в текстовом редакторе Microsoft Word и оформленную в соответствии с требованиями: формат А4, шрифт – 14 Times New Roman, интервал – полуторный; поля: левое – 30 мм; верхнее и нижнее – 20 мм; правое – 15 мм.

Не рекомендуется использовать нумерацию страниц и автоматическую расстановку переносов.

Формулы помещаются в текст с использованием редактора формул Microsoft Equation со следующими установками: обычный 14 пт; крупный индекс 9 пт; мелкий индекс 7 пт; крупный символ 18 пт; мелкий символ 12 пт.

Рисунки должны иметь четкое изображение и быть выдержаны, как правило, в черно-белой гамме.

Рисунки и таблицы должны быть пронумерованы и иметь названия; на них должны быть ссылки в тексте.

Таблицы являются частью текста и не должны создаваться как графические объекты.

Обязательным является указание УДК.

Список источников приводится в конце статьи в алфавитном порядке в соответствии с ГОСТ 7.1-2003 «Библиографическая запись. Библиографическое описание. Общие требования и правила составления».

Статья должна носить по преимуществу аналитический, а не описательный характер. В ней должен быть четкое отражение авторский подход к решению исследуемой проблемы.

2. **Аннотацию** (2-3 предложения) на русском и английском языках.

3. **Ключевые слова** на русском и английском языках.

4. **Сведения об авторе** (на русском и английском языках): ФИО полностью, ученая степень, ученое звание, место работы, должность, контактный телефон, адрес электронной почты, адрес для пересылки журнала.

Рукописи всех статей, поступивших в журнал, проходят через институт рецензирования. Максимальный срок рецензирования – от даты поступления до вынесения решения – составляет 1 месяц.

Плата с авторов за рецензирование статей не взимается. Плата за публикацию взимается в случае положительной рецензии.

Плата с аспирантов за рецензирование и публикацию статей (без соавторов) не взимается.

Авторы имеют право использовать все материалы в их последующих публикациях при условии, что будет сделана ссылка на публикацию в журнале «Современная экономика: проблемы и решения».

Материалы, не соответствующие указанным требованиям, рассматриваться не будут.