УДК 336.748.12 : 339.372.4
doi:10.18720/SPBPU/2/id23-96

Ласкин Михаил Борисович [1]
ст. научн. сотрудник, канд. физ.-мат. наук, доцент;
Пупенцова Светлана Валентиновна [2],
доцент, канд. экон. наук, доцент;
Докторов Дмитрий Всеволодович [3],
директор Проектного института

АНАЛИЗ ДАННЫХ КАССОВЫХ АППАРАТОВ В РИТЕЙЛЕ И ОЦЕНКА ИНФЛЯЦИИ

[1] Россия, Санкт-Петербург, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, laskinmb@yahoo.com
[2] Россия, Санкт-Петербург, Санкт-Петербургский политехнический университет Петра Великого, pupentsova_sv@spbstu.ru
[3] Россия, Санкт-Петербург, АО «Прикладная химия», di@giph.su

Аннотация. В работе исследуются общие вопросы оценки уровня инфляции по рыночным данным сетевого ритейла на основе двумерных распределений случайных величин. Показано, что в одной и той же группе товаров имеется диапазон с растущими ценами и диапазон, в котором цены снижаются. Используются системный и сравнительный анализ, статистические методы корреляционно-регрессионного анализа, индексные методы оценки инфляции. В качестве примера приведены данные кассовых аппаратов одного из продовольственных дискаунтеров в Санкт-Петербурге. Для объекта исследования авторами модифицированы индексы Ласпейреса и Пааше.

Ключевые слова: инфляция, уровень инфляции, дефлятор, стохастическая модель ценообразования, логарифмически нормальный закон распределения.

Mikhail B. Laskin [1],
Associate Professor, Ph.D. in Physics and Mathematics;
Svetlana V. Pupentsova [2],
Associate Professor, Candidate of Economic Sciences;
Dmitry V. Doctorov [3],
Director of the Project Institute

INFLATION RATE DEFINITION IN INVESTMENT AND CONSTRUCTION SECTOR OF THE ECONOMY

[1] St. Petersburg Federal Research Center of the Russian Academy of Sciences,
St. Petersburg, Russia, laskinmb@yahoo.com;
[2] St. Petersburg Polytechnic University of Peter the Great,
St. Petersburg, Russia, pupentsova_sv@spbstu.ru;
[3] JSC "Applied Chemistry", St. Petersburg, Russia, di@giph.su

Abstract. This paper examines the general issues of estimating the level of inflation from the market data of chain retailing on the basis of bivariate distributions of random variables. It is shown that in the same group of goods there is a range with increasing prices and a range in which prices are decreasing. System and comparative analysis, statistical methods of correlation and regression analysis, index methods of inflation estimation are used. As an example, the data of cash registers of one of the food discounters in St. Petersburg are given. The authors modified the E. Laspeyres and H. Paasche indices for the object of the study.

Keywords: inflation, inflation rate, deflator, stochastic pricing model, lognormal distribution.

Показатели инфляции активно используются как для оценки текущей ситуации, так и при макроэкономических исследованиях, прогнозировании и стратегическом планировании, поэтому нуждаются в постоянном мониторинге [1]. Для оценки инфляции используется индексные методы, описание которых представлено в [2, 3]. Ранее в работе [4] авторы провели сравнительный анализ цен и предложили алгоритм расчета инфляции для инвестиционно-строительного сектора экономики с использованием двумерных распределений. Цель данного исследования – модифицировать известные индексы Ласпейреса и Пааше для оценки уровня инфляции по рыночным данным сетевого ритейла на основе двумерных распределений случайных величин.

Для приведенного в работе примера использованы реальные данные одного из магазинов торговой сети "SPAR", работавшего в Санкт-Петербурге до 2020 года, по адресу: Лиговский проспект, д. 130. Для анализа использованы данные о ценах и объемах продажах товаров, входящих в ассортиментную матрицу указанного предприятия в марте 2018 и марте 2019 года. Общий объем списка товаров 2534.

Пусть V_1 — цена единицы товара в марте 2018 года, V_2 — цена единицы товара в марте 2019 года. Рассмотрим двумерную величину (V_1, V_2). Большими буквами будем обозначать случайные величины V_1, V_2, малыми v_1, v_2 — их значения. На рисунке 1 слева показано облако рассеяния в координатах «логарифм цены за единицу товара в марте 2018 года — логарифм цены за единицу товара в марте 2019 года». На рисунке 1 слева черная линия показывает линейный тренд для облака рассеяния логарифмов цен марта 2018 и марта 2019 года. Для изучаемого объема данных уравнение линейного тренда выглядит так:

$$\log(v_2) = 0.963 \bullet \log(v) + 0.201, \qquad (1)$$

откуда получаем оценку зависимости цены марта 2019 года от цены марта 2018 года:

$$v_2 = v_1^{0.963} \bullet e^{0.201} = 1.223 \bullet v_1^{0.963}, \qquad (2)$$

и оценку коэффициента удорожания единицы товара, в зависимости от его цены в марте 2018 года:

$$K = 1.223 \bullet v_1^{0.963-1}. \qquad (3).$$

Отметим, что более наглядными являются график оценки коэффициента удорожания K, рассчитанного по формуле (3), представленный справа на рисунке 1.

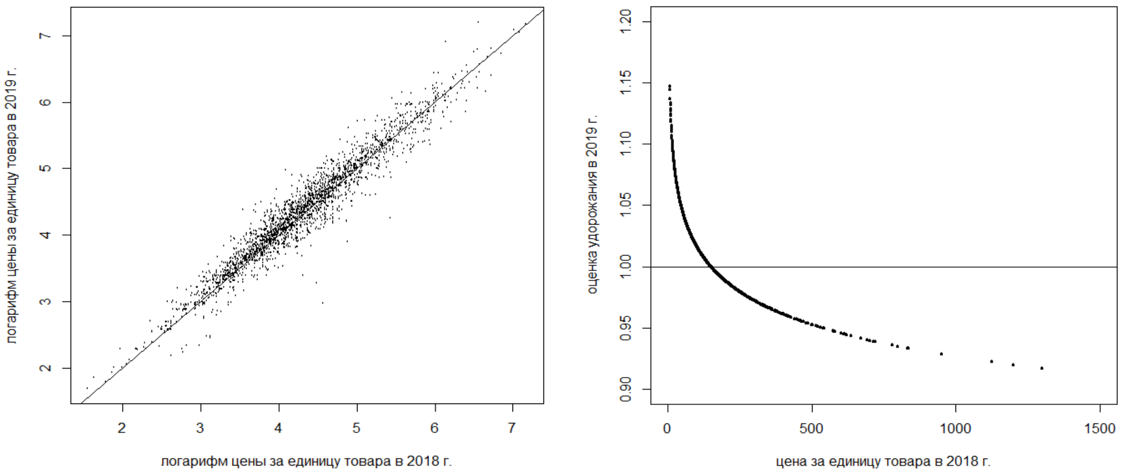

Рис. 1. Результаты анализа цен за 2018 и 2019 гг.
Слева: соотношение логарифмов цен товаров продуктовой группы в 2018 и 2019 г.
Справа: Оценка коэффициента удорожания K, соответствующая формуле (3)

Таким образом, оценка коэффициента удорожания, построенная по линейному тренду (формулы (1)–(3)) в нижнем диапазоне указывает на повышение цен, в верхнем — на снижение цен.

Принимая во внимание известное утверждение о том, что случайный вектор (V_1, V_2) нормален тогда и только тогда, когда нормальна любая линейная комбинация его компонент, авторы применили к выборке бутстреп тест [5] на проверку нормальности случайной линейной комбинации логарифмов компонент V_1, V_2.

Тестирование проводилось в следующем порядке: сначала генерируются два случайных числа по равномерному закону распределения на единичном отрезке, в сумме, равные единице; после составляется линейная комбинация логарифмов компонент V_1, V_2, линейная комбинация проверяется на нормальность тестом Колмогорова-Смирнова, результат (p-value) записывается в массив, затем все действия повторяются достаточно большое количество раз. Так, из 10 000 случайных линейных комбинаций компонент V_1, V_2 не встретилось ни одной с p-value меньше критического значения 0,05. В таких условиях мы считаем возможным рассмотреть оценку цены 2019 года по условным распределениям, вытекающим из совместной логарифмической нормальности компонент V_1, V_2.

При заданном значении переменной $V_1 = v_1$ условная мода случайной величины V_2, при условии $V_1 = v_1$ равна

$$Mode(V_2|V_1 = v_1) = \exp(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(\ln(v_1) - \mu_1) - \sigma_2^2(1 - p^2)), \quad (4)$$

условная медиана случайной величины V_2, при условии V_1=v_1 равна

$$Median(V_2|V_1 = v_1) = \exp(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(\ln(v_1) - \mu_1)), \quad (5)$$

условное математическое ожидание случайной величины V_2, при условии $V_1 = v_1$ равно

$$E(V_2|V_1 = v_1) = \exp(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(\ln(v_1) - \mu_1) + \frac{1}{2}\sigma_2^2(1 - \rho^2)). \quad (6)$$

Формулы (4), (5), (6) могут быть записаны в виде

$$Mode(V_2|V_1 = v_1) = C_{mode} \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}}, \quad (7)$$

$$Median(V_2|V_1 = v_1) = C_{median} \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}}, \quad (8)$$

$$E(V_1|v_1) = C_E \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}}, \quad (9)$$

где

$$C_{mode} = \exp(\mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1 - \sigma_2^2(1 - \rho^2)), \quad (10)$$

$$C_{median} = \exp(\mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1), \quad (11)$$

$$C_E = \exp(\mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1 + \frac{1}{2}\bullet\sigma_2^2(1 - \rho^2)). \quad (12)$$

Для исходных данных с учетом полученных средних значений, стандартных отклонений и коэффициента корреляции получаем:

$$Mode(V_2|V_1 = v_1) = 1.173 \bullet v_1^{0.963}, \tag{13}$$

$$Median(V_2|V_1 = v_1) = 1.222 \bullet v_1^{0.963}, \tag{14}$$

$$E(V_1|V_2 = v_1) = 1.248 \bullet v_1^{0.963}. \tag{15}$$

Медианная оценка по формуле (14), ожидаема, практически совпадает с оценкой по формуле (2), так как формула (5) и результат потенцирования в формуле (2) — это по существу одна и та же формула, если принять гипотезу о совместном логарифмически нормальном распределении величин V_1 и V_2.

Для коэффициента удорожания получаем три оценки (по моде, по медиане, по математическому ожиданию):

$$K_{Mode(v_1)} = \frac{Mode(V_2|V_1 = v_1)}{v_1} = C_{mode} \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}-1}, \tag{16}$$

$$K_{Median(v_1)} = \frac{Median(V_2|V_1 = v_1)}{v_1} = C_{median} \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}-1}, \tag{17}$$

$$K_{E(v_1)} = \frac{E(V_2|V_1 = v_1)}{v_1} = C_E \bullet v_1^{\rho\frac{\sigma_2}{\sigma_1}-1}, \tag{18}$$

На рисунке 2 показаны соответствующие линии, отражающие оценки коэффициента удорожания в зависимости от цены единицы товара в 2018 году.

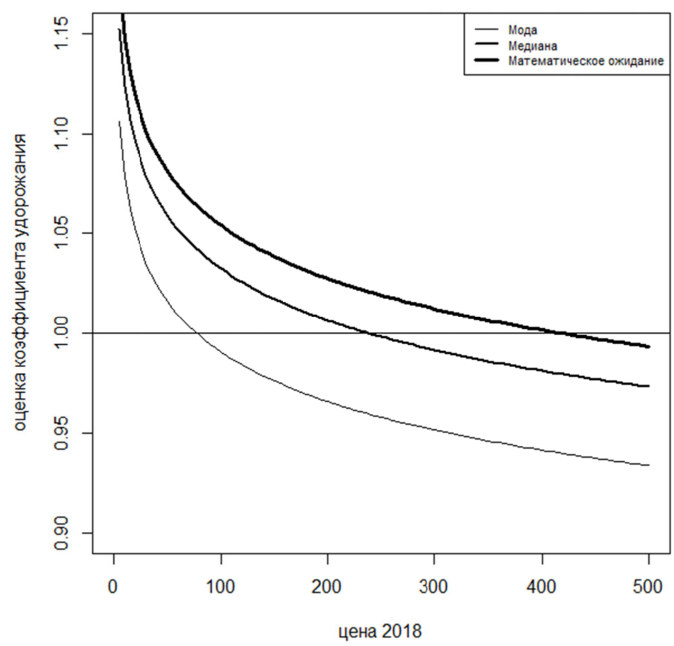

Рис. 2 Линии, отражающие оценки коэффициента удорожания в зависимости от цены единицы товара в 2018 году, полученные по моде, медиане и математическому ожиданию условных распределений величины V_2 в зависимости от величины V_1

Отметим, что основные продажи дискаунтера, предоставившего данные (в единицах товара), сосредоточены в диапазоне цен до 100 рублей за единицу товара. Обозначим $B_1(v)$ — количество продаж товаров с ценой v в марте *2018* года, $B_2(v)$ — количество продаж товаров с ценой v в марте *2019* года. Пусть A — максимальная цена единицы товара в 2018 и 2019 годах. Тогда индекс Ласпейреса рассчитывается как

$$I_L = \frac{C \bullet \int_0^A v_1^{\rho^{\frac{\sigma_2}{\sigma_1}}} B_1(v_1) dv_1}{\int_0^A v_1 B_1(v_1) dv_1}, \tag{19}$$

индекс Пааше рассчитывается как

$$I_p = \frac{C \bullet \int_0^A v_1^{\rho^{\frac{\sigma_2}{\sigma_1}}} B_2(v_1) dv_1}{\int_0^A v_1 B_2(v_1) dv_1}, \tag{20}$$

где C может принимать значения по формулам (10) – (12), в зависимости от того, какая из формул (4) – (6) применяется для оценки цены в периоде сравнения.

Для изучаемых исходных данных расчеты дают результаты, представленные в таблице 1.

Таблица 1

Расчет индексов Ласпейреса и Пааше для двух выборок

Показатель	Выборка без ограничения цены за единицу до 100 руб.		Выборка с ограничением цены за единицу до 100 руб.	
	индекс Ласпейреса, I_L	индекс Пааше, I_p	индекс Ласпейреса, I_L	индекс Пааше, I_p
Оценка по моде	0,9950	0,9945	1,0145	1,0144
Оценка по медиане	1,0369	1,0364	1,0573	1,0571
Оценка по математическому ожиданию	1,0585	1,0580	1,0793	1,0791

Отметим, что для выборки без ограничения цены за единицу до 100 руб. существенной разницы между оценками по индексу Ласпейреса и индексу Пааше нет. Так как бóльшая часть продаж у исследуемого дискаунтера сосредоточена среди товаров с диапазоном цены за единицу до 100 рублей, то результаты для выборки с заданным ограничением указывают на более высокий уровень инфляции.

Таким образом, оценка инфляции существенно зависит от того, по какой группе такая оценка проводится. Обычно, инфляция как мера повышения общего уровня цен товаров и услуг, рассчитывается по некоторой базисной потребительской корзине [6]. Однако даже такое небольшое исследование показывает, что в одной и той же группе товаров (продовольственные товары, позиционируемые дискаунтером как недорогие) имеется диапазон, в котором цены растут и диапазон, в котором цены снижаются. Также, можно заметить, что инфляция как мера повышения товаров и услуг может меняться в зависимости от местоположения предприятия торговли, его ценовой политики, покупательской способности основных клиентов. В данном примере можно заметить, что для покупателей, ориентированных на как можно больше низкую цену, годовое изменение цен может оказаться выше, чем у остальных.

Отметим, что данные сетевого ритейла являются важным источником информации для оценки общих, региональных, секторальных экономических показателей и уровня инфляции, в частности. Как можно видеть, в общей тенденции повышения цен можно выделить группы товаров, регионы, потребительские группы, ценовые диапазоны, которые могут иметь различную динамику роста или снижения цен. Дальнейшие исследования в этом направлении могут быть связаны с изучением распределений цен на один и тот же товар внутри периодов, использованных для сравнения (в нашем примере это один и тот же месяц 2018 и 2019 года) цены внутри таких периодов часто меняются в зависимости от политики поставщиков, торговых предприятий, валютных курсов и т. д. Важным остается изучение распределений в объемах продаж, которые пока не удалось удовлетворительно приблизить каким-либо теоретическим распределением.

Список литературы

1. Сапова А.К. Сравнительный анализ показателей базовой инфляции для России // Статистика и Экономика. – 2016. – № 5. – С. 63–71.

2. Козлова М.А. от Ласпейреса до Торнквиста: эмпирическая оценка индекса потребительских цен в России. Учет и статистика. – 2020. – № 2 (58). – С. 69–77.

3. Иорданова В.Г., Шапор М.А., Берцинская В.А., Захарова М.А. Тенденции развития внешней торговли России и Китая // Российский внешнеэкономический вестник. – 2021. – № 8. – С. 62–75.

4. Ласкин М.Б., Пупенцова С.В. Определение темпов инфляции в инвестиционно-строительном секторе экономики. Статистика и Экономика. – 2018. – Т. 15. № 3. – С. 14–22.

5. Bradley Efron. Bootstrap Methods: Another Look at the Jackknife (англ.) // Annals of Statistics. – 1979. – Vol. 7, no. 1. – Pp. 1–26. – ISSN 0090-5364. – doi:10.1214/aos/1176344552.

6. Климин А.И., Тихонов Д.В., Трыков А.В. Эластичность спроса при вычислении оптимальной цены // В сборнике: Фундаментальные и прикладные исследования в области управления, экономики и торговли. сборник трудов научной и учебно-практической конференции: в 3 частях. – 2017. – С. 71–76.