ВОПРОСЫ ОЦЕНКИ И НОРМИРОВАНИЯ КОЭФФИЦИЕНТА ВАРИАЦИИ

© 2017 Прокофьев Владимир Анатольевич
доктор экономических наук, профессор
© 2017 Толмачев Михаил Николаевич
доктор экономических наук, доцент, заведующий кафедрой статистики
© 2017 Головко Мария Владимировна
Саратовский социально-экономический институт (филиал)
Российского экономического университета имени Г.В. Плеханова
410003, г. Саратов, ул. Радищева, д. 89
E-mail: prokofiev@ssea.runnet.ru, tolmachev-mike@yandex.ru, golovmar@yandex.ru

Рассмотрены проблемы, возникающие при характеристике статистических совокупностей по степени однородности. Изучение характера и степени вариации у отдельных единиц совокупности является важнейшим вопросом любого исследования и позволяет судить об однородности совокупности, устойчивости значений признака, типичности средней, взаимосвязи между признаками. С целью усовершенствования статистического инструментария представлены предложения по нормированию и шкалированию коэффициентов вариации.

Ключевые слова: вариация, дисперсия, нормирование коэффициента вариации, ряды распределения, признак.

Изучение вариации в социально-экономических исследованиях имеет важное значение. Оно дает возможность определить влияние на анализируемый признак других варьирующих признаков. Кроме того, во многих монографических статистических изданиях, посвященных анализу социально-экономического развития административно-территориальных образований России с помощью коэффициента вариации, оцениваются и сравниваются в динамике, в межтерриториальном аспекте степени вариации важных социально-экономических индикаторов. При этом зачастую дают краткую и весьма приближенную характеристику степени однородности совокупностей, групп по исследуемым признакам, основываясь только на двух ненормированных градациях: меньше 33,3 % для однородных групп и более 33,3 % для неоднородных[1].

Если первая градация имеет верхнюю и нижнюю границы, то вторая - верхнего предела не имеет, и по приведенным в том или ином тексте оценкам коэффициентов вариации 50 % или 150 % трудно судить о неоднородности совокупности по определенному признаку в конкретном регионе: какой она должна или желаема быть (оптимально, по прогнозу, по плану); на сколько ее необходимо снизить или дифференцировать по группам; какие конкретные рекомендации и предложения предоставить органам законодательной или исполнительной власти.

Считаем целесообразным в пользу совершенствования статистического инструментария представить ряд предложений по нормированию и шкалированию коэффициентов вариации.

Изложим два возможных подхода к решению проблемы нормирования коэффициента вариации.

В качестве первого из них правомерным представляется установление соответствия конструкции коэффициента вариации его содержанию как средней величины относительных отклонений вариант от их средней величины:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} = \frac{1}{n} \cdot \sum \frac{\left|x - \bar{x}\right|}{\bar{x}} = \frac{1}{n} \cdot \sum \left|\frac{x}{\bar{x}} - 1\right|,$$

где x - значения признака;

$\bar{x}$ - средняя величина;

n - объем совокупности.

При этом сумма отрицательных отклонений по абсолютной величине всегда совпадает с суммой (уравновешивает сумму) положительных отклонений, но не совпадает с последней по числу отклонений.

Поэтому

$$\frac{1}{n^{+}} \cdot \sum^{+} \left(\frac{x}{\bar{x}} - 1\right) = -\frac{1}{n^{-}} \cdot \sum^{-} \left(\frac{x}{\bar{x}} - 1\right)$$

или

$$-V_{\bar{d}}^{-} = V_{\bar{d}}^{+} = \frac{1}{2} V_{\bar{d}}, \; 0 \le V_{\bar{d}}^{+} < 1,$$

где n^{+} (n^{-}) - число положительных (отрицательных) отклонений.

Таким образом, величина

$$V_{\bar{d}}^{+} = \frac{\sum\left|x-\bar{x}\right|}{2n\bar{x}} = \frac{\bar{d}}{2\bar{x}}$$

может трактоваться как нормированный коэффициент вариации на базе среднего линейного отклонения.

Иной подход к нормированию коэффициента вариации мотивируется исследованием профессора М.М. Юзбашева в определении максимально возможных значений основных показателей вариации в зависимости от их конструкции, содержания и объема совокупности[2].

Так, максимальной будет величина абсолютного отклонения, когда общий объем признака сосредоточен у одной единицы совокупности

$$\max\left|x-\bar{x}\right| = \left|\sum x-\bar{x}\right| = \left|\bar{x}n-\bar{x}\right| = \bar{x}(n-1)$$

и сумма остальных (n - 1) отклонений составляет

$$(n-1)\cdot\left|0-\bar{x}\right| = \bar{x}\cdot(n-1).$$

Следовательно,

$$\max\bar{d} = \frac{2\bar{x}\cdot(n-1)}{n}$$

и

$$\max V_{\bar{d}} = \frac{2(n-1)}{n} = \lim_{n\to\infty} \max V_{\bar{d}} = 2.$$

Подобно тому как генеральная дисперсия больше выборочной в $\frac{n}{n-1}$ раз[3], генеральная величина $V_{\bar{d}} = 2$ при $n \to \infty$ будет в $\frac{n}{n-1}$ раз больше выборочной величины $\max V_{\bar{d}} = 2\cdot\frac{n-1}{n}$ при малых n.

Значит, для оценки $V_{\bar{d}}$ в нормированной шкале [0; 1) можно принять величину

$$V'_{\bar{d}} = \frac{V_{\bar{d}}}{2}\cdot\frac{n}{n-1} \approx \frac{V_{\bar{d}}}{2}.$$

Как видим, два разных подхода демонстрируют и уточняют один и тот же результат нормирования коэффициента вариации на базе среднего линейного отклонения.

В отличие от среднего линейного отклонения природа среднего квадратического отклонения и дисперсии как среднего квадрата отклонений не отвечает прямо на вопрос оценки среднего расстояния вариант от их средней величины.

Конструкция дисперсии и среднего квадратического отклонения придает существенное усиление приоритетности больших отклонений в образовании среднего линейного отклонения с одновременным преуменьшением роли малых отклонений.

Действительно:

$$\sigma^2 = \frac{\sum(x-\bar{x})^2}{n} = \frac{\sum\left|x-\bar{x}\right|\cdot\left|x-\bar{x}\right|}{n} =$$

$$= \frac{\sum\left|x-\bar{x}\right|\cdot\left|x-\bar{x}\right|}{\sum\left|x-\bar{x}\right|}\cdot\frac{\sum\left|x-\bar{x}\right|}{n} =$$

$$= \frac{\sum\left|x-\bar{x}\right|\cdot m}{\sum m}\cdot\bar{d} = \overline{d_m}\cdot\bar{d},$$

где $m = \left|x-\bar{x}\right|$, являясь абсолютными отклонениями, одновременно выполняет роль весов абсолютных отклонений $\left|x-\bar{x}\right|$, т.е. сами себя взвешивают.

Среднее квадратическое отклонение, как и дисперсия, может быть сколь угодно большим. Его максимальная величина, как доказал профессор М.М. Юзбашев[4], для совокупности из n единиц составляет $\max\sigma = \bar{x}\cdot\sqrt{n-1}$. Соответственно, максимальная величина коэффициента вариации (V_σ) будет равна $\max V_\sigma = \sqrt{n-1}$, а $\lim\limits_{n\to\infty} \max V_\sigma = \infty$. Однако коэффициент вариации на базе среднего квадратического отклонения так же, как и на базе среднего линейного отклонения, может быть нормирован:

$$V'_\sigma = \frac{V_\sigma}{\sqrt{n-1}}.$$

По аналогии с правилом “трех сигм” нормального распределения[5] выделим шесть равных градаций между минимальной (0,0 %) и максимальной (100,0 %) величинами шкалы нормированных коэффициентов вариации по три слева и справа от центра (50,0 %). При этом поименуем дифференциацию степени варьирования схожим описанием тесноты корреляционной связи в шкале Чеддока (табл 1).

Таблица 1. **Шкала нормированного коэффициента вариации, %**

Значение	Степень вариации
0,0 - 16,7	Незначительная
16,7 - 33,3	Слабая
33,3 - 50,0	Умеренная
50,0 - 66,7	Заметная
66,7 - 83,3	Значительная
83,3 - 100,0	Высокая

Таблица 2. **Расчет нормированного коэффициента вариации**

№ п/п	Показатель	Совокупность		
		1	2	3
1	x_i	3	5	0
		5	10	0
		6	-	0
		10	-	10
2	n	4	2	4
3	$\bar{x}$	6,0	7,5	2,5
4	$\sum x$	24,0	15,0	10,0
5	$\sum^{+}(x-\bar{x})$	4,0	2,5	7,5
6	$\sum^{-}(x-\bar{x})$	-4,0	-2,5	-7,5
7	$\sum(x-\bar{x})$	8,0	5,0	15,0
8	$\bar{d}=\frac{\sum\lvert x-\bar{x}\rvert}{n}$	2,0	2,5	3,8
9	$V_{\bar{d}}=\frac{\bar{d}}{\bar{x}}$, %	33,3	33,3	150,0
10	$\max V_{\bar{d}}=2\cdot\frac{n-1}{n}$, %	150,0	100,0	150,0
11	$V'_{\bar{d}}=\frac{V_{\bar{d}}}{\max V_{\bar{d}}}$, %	22,2	33,3	100,0
12	$\sigma^2=\overline{x^2}-\bar{x}^2$	6,5	6,3	18,8
13	$\sigma=\sqrt{\sigma^2}$	2,6	2,5	4,3
14	$V_\sigma=\frac{\sigma}{\bar{x}}$, %	42,5	33,3	173,2
15	$\max V_\sigma=\sqrt{n-1}$, %	173,2	100,0	173,2
16	$V'_\sigma=\frac{V_\sigma}{\sqrt{n-1}}$, %	24,5	33,3	100,0

Рассмотрим условный пример, демонстрирующий однозначность (схожесть) выводов по величине обоих коэффициентов вариации ($V_{\bar{d}}$, V_σ) в градациях нормированной шкалы для трех разных ситуаций вариации признака (табл. 2).

Наблюдается схожесть оценок вариации $V_{\bar{d}}$ и V_σ по стр. 9 и 14 (ненормированных) с существенным превышением V_σ ($V_\sigma \geq V_{\bar{d}}$) по первой и второй совокупностям; схожесть оценок $V'_{\bar{d}}$ и V'_σ в градациях нормированной шкалы по стр. 11 и 16 с незначительным превышением V'_σ ($V'_\sigma \geq V'_{\bar{d}}$) по первой совокупности в градации слабой вариации и тождественность оценок по второй и третьей совокупностям в градациях слабой и высокой вариации.

---

[1] См.: Теория статистики : учебник / под ред. Р.А. Шмойловой. 3-е изд., перераб. Москва, 2001. С. 187; Общая теория статистики : учебник / Т.В. Рябушкин [и др.]. Москва, 1981. С. 107; Социально-экономическая статистика : учеб. для вузов / под ред. Б.И. Башкатова. Москва, 2002. С. 66.

[2] *Елисеева И.И., Юзбашев М.М.* Общая теория статистики : учебник / под ред. И.И. Елисеевой. Москва, 1995. С. 100-103.

[3] Там же. С. 142.

[4] Там же. С. 101.

[5] *Гмурман В.Е.* Теория вероятностей и математическая статистика : учеб. пособие для вузов. Москва, 1998. С. 134-135.

Поступила в редакцию 04.09.2017 г.