

НЕКОТОРЫЕ ОСОБЕННОСТИ ПРИМЕНЕНИЯ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА В ИССЛЕДОВАНИИ ПОКАЗАТЕЛЕЙ ДЕЯТЕЛЬНОСТИ ПРЕДПРИЯТИЯ

Введение. Методы корреляционно-регрессионного анализа для изучения взаимосвязей различных явлений являются основными и хорошо разработанными. Для построения статистически достоверной регрессионной модели необходимо иметь достаточно большое количество исходных пространственных данных для множества объектов. Это могут быть данные по множеству предприятий одной отрасли собранных в одно и тоже время для изучения взаимосвязи таких факторов как прибыль, затраты, выручка, заработная плата, производительность труда и других. Но такие же задачи возникают и в рамках изучения деятельности одного предприятия. В этом случае исходные данные по интересующим нас показателям могут быть собраны на одном предприятии как годовые показатели за ряд лет. Напрямую регрессионный анализ в этом случае применять нельзя, т.к. сами показатели изменяются во времени и мы имеем дело с взаимосвязанными рядами динамики.

Основная часть. Применение корреляции в динамических рядах имеет ряд особенностей, недоучет которых не позволяет получить правильной оценки взаимосвязи между рядами динамики, которые, в свою очередь, рассматриваются как результирующий и факторный признаки [1].

Для построения адекватного уравнения регрессии, характеризующего взаимосвязь интересующих показателей необходимо сначала исключить трендовые составляющие из исходных динамических рядов, а затем строить уравнение регрессии на остатках. В результате направление и характер взаимосвязи может сильно отличаться от первоначального. Покажем это на примере. Допустим, изучается взаимосвязь объема реализованной продукции (выручка без НДС) и средней заработной платы работников одного предприятия в условных денежных единицах за тринадцать последовательных лет. По исходным данным построим корреляционное поле и оценим параметры линейного уравнения методом наименьших квадратов (МНК) (рисунок 1).

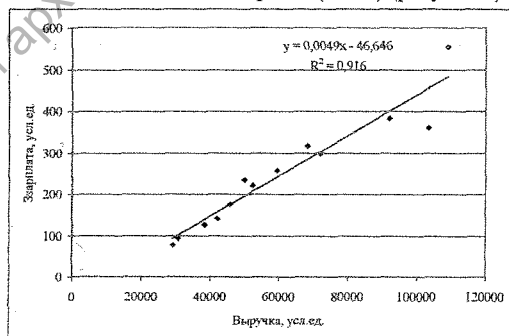


Рис. 1. Взаимосвязь показателей объема реализованной продукции и средней заработной платы

Анализируя полученное уравнение регрессии можно сказать, что оно статистически достоверное по критерию Фишера ($F = 111,5, p < 0,05$) и объясняет около 90% эмпирических данных. Но при исследовании остатков данной модели обнаруживается их автокорреляция, что не является признаком адекватной модели. Наличие автокорреляции положительной или отрицательной свидетельствует о неправильной спецификации модели [2, 3]. Зависимость между переменными может носить на самом деле нелинейный характер, либо имеется неучтенный в уравнении неизвестный фактор.

Данное явление оценивается с помощью критерия Дарбина – Уотсона, который рассчитывается:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

где $e_t = Y_t - \hat{Y}_t$ – остатки, рассчитанные по уравнению регрессии;

Y_t – исходные данные результативного признака, в нашем случае средняя заработная плата;

\hat{Y}_t – теоретические значения результативного признака, полученные по уравнению регрессии.

Не обращаясь к таблице критических точек Дарбина – Уотсона, можно пользоваться «грубым» правилом и считать, что автокорреляция остатков отсутствует, если $1,5 < DW < 2,5$. Для нашего уравнения регрессии значение $DW = 1,35$, что свидетельствует о наличии положительной автокорреляции.

Анализируя динамику исходных показателей можно увидеть их сильную статистически значимую зависимость от времени имеющую практически одинаковый характер (рисунки 2 и 3).

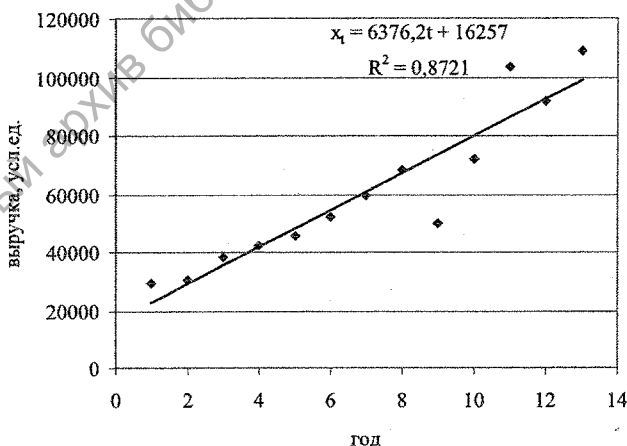


Рис. 2. Динамика показателя объема реализованной продукции

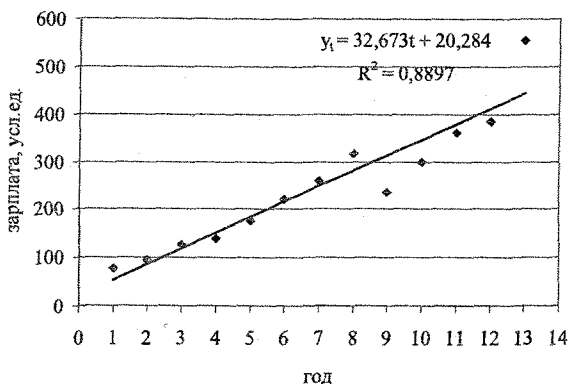


Рис. 3. Динамика показателя среднемесячной заработной платы

Данным обстоятельством нельзя пренебрегать, так как время в нашем случае и является тем самым неучтенным фактором. Учесть его можно исключив трендовую составляющую из каждого ряда, рассчитав остатки, а затем, построить новое уравнение регрессии для полученных значений остатков.

Таким образом, имея уравнение тренда для объема реализованной продукции ($x_t = 6376,2t + 16257$) рассчитаем остатки:

$$e_x = x - x_t$$

По уравнению тренда для заработной платы ($y_t = 32,673t + 20,284$) также рассчитаем остатки:

$$e_y = y - y_t$$

На полученных значениях остатков e_x и e_y построим новое уравнение регрессии (рисунок 4):

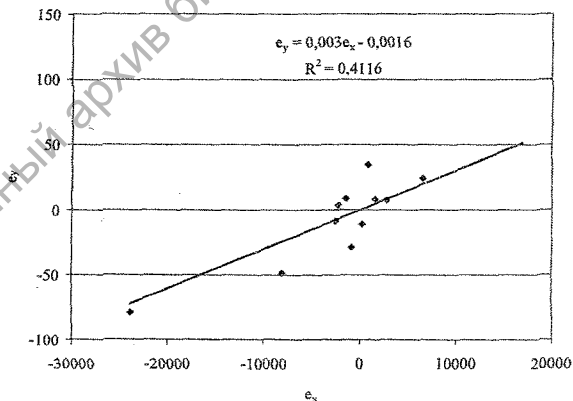


Рис. 4. Взаимосвязь величин средней заработной платы и объема реализованной продукции с исключенными трендовыми составляющими

Это уравнение также является статистически значимым ($F = 7.5$, $p < 0.05$) коэффициент детерминации $R^2 = 0.411$, говорит о том, что около 40% вариации заработной платы объясняется вариацией объема реализованной продукции. Кроме того данная модель характеризуется отсутствием автокорреляции остатков, величина критерия Дарбина – Уотсона $DW = 2.28$, что является признаком более адекватной модели.

Заключение. Данный пример иллюстрирует подход, который позволяет в отсутствии пространственных данных и располагая только данными наблюдения по годам для отдельно взятого предприятия выявить истинные характеристики как парных так и множественных взаимосвязей методами корреляционно-регрессионного анализа не нарушая предпосылки МНК [2].

Список использованных источников

1. Общая теория статистики : учебник / под ред. чл.-кор. РАН И.И. Елисеевой. – 4-е изд., перераб. и доп. – М. : Финансы и статистика, 202. – 480 с. : ил.
2. Бородич, С.А. Эконометрика / С.А. Бородич. – Минск : ООО Новое знание. – 2001. – 407 с. : ил.
3. Магнус, Я.Р. Эконометрика. Начальный курс / Я.Р. Магнус, П.К. Катъшев. – М. : Дело, 2000.