PERIODICA
Journal of Modern Philosophy, Social Sciences and Humanities
ISSN NO:2720-4030 Volume 42 May 2025

Математико-статистический подход к изучению однородности экономической информации при использовании корреляционно-регрессионного метода

Сирожиддинов Икромиддин Кутбиддинович, канд.экон. наук, проф.
Ботирова Рахимахан Абдужаббаровна, канд. экон. наук, доц.
Наманганский государственнқй технический университет

ABSTRACT

в статье исследован математико-статистический подход к изучению однородности экономической информации при использовании корреляционно-регрессионного метода на конкретном примере.

ARTICLE INFO

Received: 26th February, 2025
Accepted: 24th March 2025

KEYWORDS:
экономические процессы, корреляция, математико-статистический подход, субъекты агробизнеса, однородность совокупности, критерии согласия, нормальное распределение.

Сущностное содержание экономических процессов проявляется в экономических показателях. Научное исследовании основывается на обработке, сравнении и анализе показателей или признаков. Общеизвестно, что при исследовании экономических процессов сравниваемые явления должны иметь общую для них сущностную основу, так как сравнение данных с различным содержанием не обеспечивает объективный характер получаемых выводов.

Выбор методов в экономическом исследовании зависит от формы проявления изучаемых закономерностей.Так, успешное применение метода корреляционного анализа связано с законом пормального распределения изучаемых случайных величин. При применении данного метода определение статистических характеристик существенности связи и построение доверительных интервалов исходят из наличия нормального распределения.

Кроме того, корреляционный анализ предполагает также однородность исследуемой совокупности по изучаемым признакам. В экономическом смысле однородность совокупности означает, что ее элементы формировались подопределящим воздействием общих для факторов. Однако, ознакомление с экономической литературой показывает, что часто при применении данного метода для анализа экономических явлений, не указывают имело ли место нормальное распределение изучаемых признаков и не приводят оценку однородности исследуемой совокупности.

Целями данного исследования являются изучение взаимосвязи между однородностью совокупности и нормальностью распределения экономических признаков, рассмотрение некоторых

аспектов использования математическо-статистических методов для определения степени приемлемости исходной экономической информации при корреляционном анализе.

Нормальное распределение имеет место, когда на изучаемую случайную величину влияет большое число факторов, воздействия которых преимущественно независимы и примерно одинаковы, а также среди этих факторов нет ни одного резко выделяющегося колебаниями своих значений или изменяемостью (колебляемость). При достаточном количестве наблюдений экономические явления в большинстве своем имеют, хотя и не строго нормальное, но все же довольно близкое к нормальному распределение.

Экономические явления формируются под влиянием причин и условий двоякого рода:

1) основных, типичных, общих для всех объектов исследуемой совокупности;

2) нетипичных, индивидуальных причин и условий, создающих специфику, индивидуальные особенности отдельных объектов совокупности.

Факторы первой группы определяют массовое состояние изучаемого экономического явления, формируют его типичный уровень и, хотя влияние этих факторов нельзя признать абсолютно одинаковым, но, как правило, вариация изучаемого признака, вызванная их влиянием, ограничена определенными рамками. Действие этой группы факторов в значительной степени обуславливает нормальное распределение величин изучаемого признака. Что касается второй группы факторов, их влияние не является одинаковым во всех единицах совокупности и вызывает отклонение от типичного уровня изучаемого признака.

Таким образом, нетипичные факторы препятствуют более полному проявлению влияния факторов первой группы, ухудшают «нормальность» распределения изучаемого признака. Исключив из исследуемой совокупности данные объекты, можно улучшить однородность совокупности и, соответственно, «нормальность» распределения изучаемого признака.

Для выявления объектов, характеристики которых нетипичны, возможно применение критериев двух видов: логический (качественный) и статистический (количественный). При применении логического критерия все объекты совокупности подвергаются подробному обследованию, призванному установить в какой степени в каждом объекте значение изучаемого признака обусловлено специфичными факторами.

Для выявления резко выделяющихся наблюдений целесообразно сначала использовать статистические критерии, а затем «выскочившие» наблюдения подвергать логическому контролю для окончательного решения вопроса о принадлежности данного объекта к исследуемой совокупности. Сущность статистических крнтериев заключается в том, что исключаются наблюдения, вероятность появления которых очень мала. Среди них наиболее простым является так называемое «правило трех сигм». Меньшее распространение получил критерий Шовене (устраняются те случаи, вероятность появления которых в два раза меньше числа наблюдений).

Исключение резко выделяющихся наблюдений может улучшить «нормальность» распределения, но еще не гарантирует соответствие эмпирического распределения нормальному и, следовательно, условиям применимости метода корреляционного анализа.

Различия между эмпирическим и нормальным распределением характеризуют показатели асимметрии и эксцесса. Асимметричность и эксцессивность кривой распределения могут быть вызваны неоднородностью совокупности. Если в какой-то группе объектов преобладает влияние нетипичных факторов, то это приведет к смещению центра распределения и вызывает асимметрию. Далее объединение различных совокупностей приводит к бимодальности или мультимодальности и тем самым ухудшает показатель эксцесса.

Таким образом, установление причин, вызывающих асимметрию и эксцессивность кривой распределения, позволяют на их основе не только улучшить сами эти показатели и, следовательно, «нормальность» распределения, но и в значительной степени обеспечивает однородность совокупности.

Необходимо заметить, что у статистиков нет единого мнения по вопросу при каких значениях асимметрии и эксцесса различия между эмпирическим и нормальным распределениями можно

признать несущественными. Так Е. А. Дмитриев указывает, что условно распределения можно считать слабо асимметричными, если |A|<0,5, средне асимметричными при 0,5<|A|1,0 и сильно асимметричными при A|>1,0[1]. Авторы учебника «Общая теория статистики» Н. М. Виноградова и другие считают асимметрию значительной при A|>0,5[2] и так далее.

Показатели асимметрии и эксцесса не дают вероятностную оценку существенности различий между эмпирическим и нормальным распределением. Такую оценку дают так называемые критерии согласия[3].

Среди критериев согласия наиболее распространенным является критерий Пирсона (χ^2 - хи-квадрат), рассчитываемый по формуле:

$$\chi^2 = \sum_{i=1}^{i=l} \frac{(n_i - n_i^1)^2}{n_i}$$

Где n_i-эмпирические частоты, $n_i^{'}$ - теоретические частоты, l- количество групп.

Как видно из формулы, при вычислении численного значения этого критерия исследуемая совокупность делится на несколько групп (интервалов) и для каждой из этих групп сравнивают теоретические и эмпирические частоты (при этом интервалы групп должны устанавливаться таким образом, чтобы в каждую группу попало не менее 5-6 объектов, а совокупность должна включать не менее 50 объектов). Если эмпирические и теоретические частоты существенно отличаются друг от друга, то гипотеза о соответствии эмпирического распределения теоретическому отклоняется (с определенной вероятностью).

Среди остальных критериев согласия отметим критерии Колмогорова (*Л - лямбда*) и Мизеса-Смирнова (nw^2 - эн омега квадрат). В отличие от критерия Пирсона критерии *Л* и nw^2 предназначены для рядов распределений, равенство любых двух членов которых исключено. Однако, исследователи пользовались этими критериями и в тех случаях, когда такое условие не выполнялось[4].

Была поставлена задача на фактическом материале провести сравнительный анализ различных критериев согласия для оценки существенности различий между эмпирическим и нормальным распределениями (таблица 1).

Таблица 1

Статистические характеристики одномерных рядов распределения субъектов агробизнеса Наманганской области по некоторым экономическим показателям

Показатели (в расчете на 1га)	Характеристики								
	- X	сигма	Коэф. вар.	Асимметрия	Эксцесс	Лямбда	nw^2	X^2	Число групп
Качество пашни (баллы)	63,68	9,64	15,2	-0,568	-0,17	0,84	0,040	8,66	6
Водообеспеченность (тыс. метр куб)	12,39	2,89	23,5	0,209	-0,752	1,24	0,306	19,59	7

[1] Дмитриев Е. А. Математическая статистика в почвоведении. -Москва, издательство Моск. Гос унив.,1972, стр. 39.
[2] Виноградова Н. М. и другие. Общая теория статистики. Учебник. – Москва, Статистика, 1968, стр. 137.
[3] Годин А. М. Статистика. Учебник для бакалавров.12-е изд., стер.—Москва, Издательско-торговая корпорация «Дашков и К°», 2020, стр. 197.

[4] Карасев А. И. Основы математической статистики. – Москва, Росвузиздат, 1962, стр. 193.

Основные средства (тыс. сумов)	1426,9	362,48	25,2	1,453	4,192	0,76	0,195	5,57	5
Производственные затраты (тыс. сумов)	1357,8	161,98	12,0	0,620	1,30	0,57	0,067	2,63	5
Оплата труда (тыс. сумов)	786,7	116,76	14,9	0,681	1,039	0,48	0,085	2,45	5
Затраты труда (чел.-часов)	1271,8	244,45	19,3	0,548	1,107	0,41	0,049	0,10	5
Количество среднегодовых работников	0.76	0,18	24,6	0,609	0,577	0,54	0,080	1,94	5
Валовая продукция (тыс. сумов)	1515,8	212,26	14,1	0,268	0,768	0,60	0,066	1,30	5

Примечание. Критические значения критериев согласия для общепринятых уровней значимости составляют: при q = 0,2 $Л = 1,08$ и nw^2 = 0,24; при q = 0,05 для 2-х степеней свободы X^2=5,99 и для 3-х степеней свободы χ^2=7,01.

В таблице приведены характеристики одномерных рядов распределения 94 субъектов агробизнеса Наманганской области по некоторым экономическим признакам.

Согласно данным этой таблицы среди перечисленных критериев согласия наиболее строгим является критерий X^2. Так, на уровне значимости q = 0,05 гипотезы о «нормальности» распределения для совокупности колхозов Наманганской области отклонялись по критерию X^2 для двух признаков, по критериям $Л$ и nw^2 – для одного признака из 8. Если сравнить критерии $Л$ и nw^2, то у них примерно равная «строгость». Так, на уровне значимости q = 0,01 гипотеза о соответствии распределения нормальному отклоняется по первому из них для двух признаков, а по второму – для одного. Но на уровне значимости q = 0,3 по обоим критериям нужно признать не соответствующими нормальному распределению 2 признаков. Заметим, что на величину X^2 может повлиять и неудачная группировка. Поэтому в случаях отклонения гипотезы о соответствии распределения нормальному по критерию X^2 целесообразна перепроверка и другими критериями согласия.

Показатели асимметрии и эксцесса также могут выступать в роли критериев оценки «нормальности» распределения: по критериям согласия распределение колхозов по обеспеченности основными производственными фондами сельскохозяйственного назначения следует признать близким к нормальному, но согласно рекомендациям о границах асимметрии, отклонение кривой эмпирического распределения от нормальной весьма существенное.

Из таблицы видим, что распределение субъектов агробизнеса области по водообеспеченности в вегетационном периоде не соответствует нормальному. С помощью вышеуказанных методов попытаемся улучшить «нормальность» распределения по данному признаку и посмотрим каки это отразится на величинах коэффициентов корреляции и регрессии, выражающих зависимость между уровнем водообеспеченности в вегетационном периоде и выходом валовой продукции в в сопоставимых ценах в расчете на 1 га орошаемой пашни. Первоначальный вид уравнения регрессии следующая: y = 1397,48 + 9,28x, где x – водообеспеченность. Наличие связи между изучаемыми явлениями статистически не подтверждается: r = 0,126; t = 1,22.

Установим причины, вызвавшие отклонение распределения аргумента от нормального. В нашем случае их два. Во-первых, высокий уровень водообеспеспеченности имеют хозяйства Мингбулакского района с очень низким качеством орошаемой пашни, представляющие собой песчаные почвы. В этих условиях выращивание сельскохозяйственных культур возможно только при

высокоинтенсивном орошении, то есть большой расход воды в этих хозяйствах есть своеобразное возмещение низкого качества пашни. Во-вторых, с вводом Большого Наманганского канала Уйчинский район имеет гарантированный и очень высокий уровень водообеспеченности.

Повышение уровня водообеспеченности, как и любого другого фактора производства, даст ожидаемый экономический эффект при соответствующем усовершенствовании технической и технологической основы производства, в противном случае ожидаемый экономический эффект может быть получен только при вовлечении в сельскохозяйственный оборот новых земель. В хозяйствах Уйчинского района выполнение этих условий не достигнуто.

Все это дает основание исключить субъекты агробизнеса Уйчинского и Мингбулакского районов (17 хозяйств) из исследуемой совокупности. После этого по признаку водообеспеченность в вегетационном периоде показатели асимметрии и эксцесса составят соответственно 0,521 и 0,0,02, а гипотеза о соответствии распределения подтверждается по критерию X^2. Зависимость между изучаемыми явлениями выражается уравнением регрессии y = 1182,66x + 27,78x. Наличие связи между изучаемыми явлениями статистически подтверждается: r = 0,30; t = 2,74.

Таким образом, улучшение «нормальности» распределения и однородности совокупности позволило выявить зависимость между изучаемыми явлениями.

Подытоживая вышеизложенное, можно заключить, что примерный алгоритм определения и улучшения состава исследуемой совокупности для проведения корреляционного анализ выглядит следующим образом:

1) устанавливаются основные, типичные причины и условия, формирующие данное экономическое явление, на этой основе определяется однородная совокупность объектов и выясняется, может ли иметь место нормальное распределение изучаемого признака;

2) при помощи критериев согласия оценивается существенность различия между эмпирическим и нормальным распределениями;

3) на основе «правила трех сигм», критерия Шовене, показателей асимметрии и эксцесса выявляются объекты, в которых значение изучаемого признака в существенной мере обусловлено влиянием нетипичных факторов;

4) объекты с резко выделяющимися значениями изучаемого признака подвергаются логическому контролю: выявляются нетипичные факторы, обусловившие чужеродность данных объектов и решается вопрос о принадлежности данного объекта к исследуемой совокупности;

5) после исключения чужеродных наблюдений еще раз оценивается существенность различий между эмпирическим и нормальным распределением: если различия не существенны, то можно приступить к проведению корреляционного анализа, в противном случае придется либо отказаться от применения данного метода, либо необходимо более детально изучить исследуемую совокупность и выявить причины, породившие несоответствие распределения нормальному.

Теперь, используя данный алгоритм, попытаемся проверить «нормальность» распределения и однородность совокупности субъектов агробизнеса Наманганской области по другим экономическим признакам. По данным таблицы распределение субъектов по качеству орошаемой пашни обладает некоторой левосторонней асимметрией и незначительным эксцессом. Гипотеза о «нормальности» распределения принимается по критериям Колмогорова и Мизеса-Смирнова. Все это дает основание признать отклонение данного распределения от нормального несущественным.

По признаку обеспеченности основными средствами гипотеза о «нормальности» распределения принимается по всем трем критериям согласия. Однако данное распределение характеризуется сильной асимметрией и значительным эксцессом. Причиной этого явилось показатели двух субъектов, не укладывающихся в интервал трех сигм. Это субъект в Янгикурганском районе и субъект в Папском районе. Изучение экономических показатели этих хозяйств показывает, что обеспеченность основными фондами субъекта Папского района превышает среднеобластной уровень примерно в два раза. Однако, значительную часть основных средств в хозяйстве составляют ирригационные сооружения, в силу чего хозяйство на 1 га пашни получает воды больше, чем соседние примерно в 1,5 раза, а материальные затраты в расчете на 1 га пашни у них одинаковы. Следовательно,

структура основных средств данного субъекта существенно отличается от аналогичного показателя в других хозяйствах, в силу чего по обеспеченности основными средствами его нельзя сравнивать с другими хозяйствами. Поэтому этот субъектисключен из дальнейшего рассмотрения.

Субъект Янгикурганского района несмотря на высокую обеспеченность основными средствами (выше среднеобластного в 2 с лишним раза) в расчете на 1 га получил продукции больше среднеобластного уровня только на 20%. Кроме того, в хозяйстве очень высокий уровень затрат труда на 1 га (среднеобластной уровень превзойден в 1,8 раза). Все это вызвано тем, что земли хозяйства мелко контурные; посевные площадей размещены на контурах размером 1-2 гектара. Это отрицательно влияет на использование техники, труда и оросительной воды. Поэтому и это хозяйство был признано чужеродным для исследуемой совокупности.

По всем остальным признакам показатели других хозяйств не выходят за пределы интервала трех сигм, а гипотеза о соответствии распределения нормальному принимается по всем критериям. Что касается показателей асимметрии и эксцесса, то они изменяются соответственно в пределах 0,27 – 0,8 и 0,58 – 1,30. Такие значения асимметрии и эксцесса можно признать несущественными.

В таблице 2 приведены характеристики одномерных рядов распределений после исключения из совокупности двух вышеназванных субъектов.

Таблица 2

Статистические характеристики одномерных рядов распределения субъектов агробизнеса Наманганской области после исключения резко выделяющихся наблюдений

Показатели (в расчете на 1га)	Характеристики								
	- X	сигма	Коэф. вар.	Асимметрия	Эксцесс	Лямбда	nw^2	X^2	Число групп
Качество пашни (баллы)	63,77	9,63	15,1	-0,576	-0,253	0,83	0,037	1,95	5
Основные средства (тыс. сумов)	1394,9	292,84	21,0	0,558	1,0	0,74	0,103	4,48	7
Производственные затраты (тыс. сумов)	1351,9	150,51	11,1	0,183	1,019	0,50	0,050	1,75	5
Оплата труда (тыс. сумов)	781,7	108,65	13,9	0257	-0,382	0,55	0,067	3,07	6
Затраты труда (чел.-часов)	1265,5	211,34	16,8	0,638	2,202	0,56	0,053	3,33	7
Количество среднегодовых работников	0.75	0,18	24,0	0,488	0,601	0,49	0,066	1,76	5
Валовая продукция (тыс. сумов)	1511,8	211,89	14,0	0,280	0,847	0,57	0,065	1,90	5

Примечание. Критическое значение χ^2 при $q = 0,05$ для 4-х степеней свободы равно 9,488.

Как показывают данные этой таблицы, по всем рассматриваемым признакам распределение субъектов агробизнеса области по большинству критериев согласия не существенно отличаются от нормального. На основе проведенного анализа можно считать, что совокупность субъектов агробизнеса области после исключения нетипичных наблюдений является качественно однородной. Эти обстоятельства позволяют сделать вывод о том, что соблюдаются условия применения методов

корреляционного анализа и аналитических группировок имеются для исследования взаимосвязей между указанными выше показателями.

Литература:

1. Дмитриев Е. А. Математическая статистика в почвоведении. - Москва, издательство Моск. Гос унив.,1972, стр. 39.
2. Виноградова Н. М. и другие. Общая теория статистики. Учебник. – Москва, Статистика, 1968, стр. 137.
3. Годин А. М. Статистика. Учебник для бакалавров.12-е изд., стер.—Москва, Издательско-торговая корпорация «Дашков и К°», 2020, стр. 197.
4. Карасев А. И. Основы математической статистики. – Москва, Росвузиздат, 1962, стр. 193.
5. Soatov N. M., Nabiyev X., Ayubjonov A. H. Statistika fanidan Darslik. Тошкент, 2019.
6. Abdulazizovich, X. U. B. (2025). INNOVASION FAOLIYATNING MOLIYAVIY TA'MINOTINI KUCHAYTIRISH ZARURLIGI VA OMILLARI TAHLILI. *Scientific Journal of Actuarial Finance and Accounting*, *5*(03), 142-146.
7. Abdulazizovich, X. U. B. (2025). MOLIYAVIY AKTIVLAR HISOBINING USLUBIY ASOSLARINI TAKOMILLASHTIRISH. *ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ*, *62*(3), 146-152.
8. Abdulazizovich, X. U. B. (2024). MOLIYAVIY AKTIVLAR TASNIFINI TAKOMILLASHTIRISH. *ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ*, *41*(2), 82-86.
9. Qutbiddinovich, S. I., & Abdulazizovich, X. U. B. (2023). GAAP ASOSIGA QO'YILGAN MOLIYAVIY HISOB KONTSEPTSIYALARI. *Interpretation and researches*, *1*(3), 42-50.
10. Abdulazizovich, K. U. (2023). POSITIVE ASPECTS OF THE CASH METHOD IN SMALL ENTERPRISES UNDER UNUSUAL CIRCUMSTANCES. *ASIA PACIFIC JOURNAL OF MARKETING & MANAGEMENT REVIEW ISSN: 2319-2836 Impact Factor: 8.071*, *12*(11), 38-47.
11. Rakhimjanovna, M. N. (2024). PRODUCT DIVERSIFICATION AS A MEANS TO ENHANCE THE ECONOMIC STABILITY OF ENTERPRISES. *Western European Journal of Medicine and Medical Science*, *2*(11), 49-55.
12. Nazira, M. (2024). KICHIK BIZNES VA TADBIRKORLIKDA MARKETINGNI QO 'LLASH TENDENSIYALARI. *Scientific Journal of Actuarial Finance and Accounting*, *4*(03), 44-50.
13. Botirova, R. A., & Sirojiddinov, I. Q. (2024). Iqtisodiyotning erkinlashtirilishi sharoitlarida kichik biznes sektori rivojlanishining imkoniyatlari. *YASHIL IQTISODIYOT VA TARAQQIYOT*, *1*(4).
14. Botirova, R. A., & Sirojiddinov, I. Q. (2023). CURRENT ISSUES IN ACCOUNTING FOR FINANCIAL INVESTMENTS. *INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE, IT, ENGINEERING AND SOCIAL SCIENCES ISSN: 2349-7793 Impact Factor: 6.876*, *17*(05), 12-15.
15. Qutbiddinovich, S. I., & Abdujabborovna, B. R. (2022). Investment processes as a key factor in the development and competitiveness of the national economy.
16. Сирожиддинов, И. К., & Исомухамедов, А. Б. (2022). Развитие экспортной направленности сельскохозяйственного производства.