

ИСПОЛЬЗОВАНИЕ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА

Суханова Е.А.

УО «Гродненский государственный аграрный университет»
г. Гродно, Республика Беларусь

Применение корреляционного анализа позволяет установить наличие или отсутствие связи между изучаемыми величинами в условиях действия большого числа факторов, ряд из которых неизвестен. Наличие или отсутствие связи определяется путем составления уравнений связи или, как их принято называть, уравнений регрессий. Форму связи (прямолинейную или криволинейную) определяют на основе предварительного анализа. О наличии корреляционной связи, а также о форме и характере кривой в первом приближении может дать представление графический анализ распределения совокупности точек, образующих корреляционное поле [2]. Корреляционная связь является статистической, условием возможности ее изучения является наличие данных по достаточно большой совокупности явлений. По отдельным явлениям можно получить совершенно превратное представление о связи признаков, ибо в каждом отдельном явлении значения признаков кроме закономерной составляющей имеют случайное отклонение (вариацию). Например, сравнивая два хозяйства, одно из которых имеет лучшее качество почв, по уровню урожайности, можно обнаружить, что урожайность выше в хозяйстве с худшими почвами. Ведь урожайность зависит от сотен факторов и при том же самом качестве почв может быть и выше, и ниже. Но если сравнивать большое число хозяйств с лучшими почвами и большое число с худшими, то средняя урожайность в первой группе окажется выше, и станет возможным измерить достаточно точно параметры корреляционной связи. Число явлений, достаточных для анализа корреляционной связи, зависит от цели анализа, требуемой точности и надежности параметров связи, от числа факторов, корреляция с которыми изучается. Обычно считают, что число наблюдений должно быть не менее чем в 5-6, а лучше - не менее чем в 10 раз больше числа факторов. Еще лучше, если число наблюдений в несколько десятков или в сотни раз больше числа факторов, тогда закон больших чисел, действуя в полную силу, обеспечивает эффективное взаимопогашение случайных отклонений от закономерного характера связи признаков [1].

Помимо этого необходима достаточно качественная однородность совокупности. Нарушение этого условия может нарушить параметры корреляции. Например, в массе зерновых хозяйств уровень продукции с гектара растет по мере концентрации площадей, т.е. он выше в крупных хозяйствах. В массе овощных и овоще-молочных хозяйств (пригородный тип) наблюдается та же прямая связь уровня продукции с размером хозяйства. Но если соединить в общую неоднородную совокупность те и другие хозяйства, то связь уровня продукции с размером площади пашни (или посевной площади) получится обратной. Причина в том, что овощные и овоще-молочные хозяйства, имея меньшую площадь, чем зерновые, производят больше продукции с гектара ввиду большей интенсивности производства в данных отраслях, чем в производстве зерна [1].

Корреляционная связь не предполагает причинной зависимости между переменными. Корреляционный анализ может использоваться для определения тесноты и направления связи. Инструментами корреляционного анализа являются разнообразные меры связи. Выбор мер (коэффициентов) связи зависит от способов измерения переменных и характера связи между ними. Корреляционно-регрессионный анализ учитывает межфакторные связи, следовательно, дает более полное измерение роли каждого фактора: прямое, непосредственное его влияние на резульативный признак; косвенное влияние фактора через его влияние на другие факторы; влияние всех факторов на резульативный признак.

ЛИТЕРАТУРА

1. Общая теория статистики: Под ред. И. И. Елисевой – Условия применения и ограничения корреляционно-регрессионного анализа – М., 2012
2. <http://www.ngpedia.ru/id313934p1.html>