ВВЕДЕНИЕ

В экономических исследованиях часто решают задачу выявления факторов, определяющих уровень и динамику экономического процесса. Такая задача чаще всего решается методами регрессионного анализа. Для достоверного отображения объективно существующих в экономике процессов необходимо выявить существенные взаимосвязи и не только выявить, но и дать им количественную оценку. Этот подход требует вскрытия причинных зависимостей. Под причинной зависимостью понимается такая связь между процессами, когда изменение одного из них является следствием изменения другого.[2]

Корреляционный анализ и регрессионный анализ являются смежными разделами математической статистики, и предназначаются для изучения по выборочным данным статистической зависимости ряда величин; некоторые из которых являются случайными. При статистической зависимости величины не связаны функционально, но как случайные величины заданы совместным распределением вероятностей.

Исследование зависимости случайных величин приводит к моделям регрессии и регрессионному анализу на базе выборочных данных. Теория вероятностей и математическая статистика представляют лишь инструмент для изучения статистической зависимости, но не ставят своей целью установление причинной связи. Представления и гипотезы о причинной связи должны быть привнесены из некоторой другой теории, которая позволяет содержательно объяснить изучаемое явление.[6]

Задачи регрессионного анализа лежат в сфере установления формы зависимости, определения функции регрессии, использования уравнения для оценки неизвестных значении зависимой переменной.

Решение названных задач опирается на соответствующие приемы, алгоритмы, показатели, применение которых дает основание говорить о статистическом изучении взаимосвязей.[11]

Целью курсовой работы является исследование регрессионного анализа в статистическом изучении взаимосвязи показателей.

Задачи курсовой работы:

- статистическое изучение взаимосвязи социально - экономических явлений и процессов;

- рассмотрение регрессионного анализа;

- исследование регрессионного анализа для изучения объекта исследования.

1. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ И ПРОЦЕССОВ

Исследование объективно существующих связей между социально-экономическими явлениями и процессами является важнейшей задачей теории статистики. В процессе статистического исследования зависимостей вскрываются причинно-следственные отношения между явлениями, что позволяет выявлять факторы (признаки), оказывающие основное влияние на вариацию изучаемых явлений и процессов. Причинно-следственные отношения - это такая связь явлений и процессов, когда изменение одного из них - причины ведет к изменению другого - следствия.

Финансово-экономические процессы представляют собой результат одновременного воздействия большого числа причин. Следовательно, при изучении этих процессов необходимо выявлять главные, основные причины, абстрагируясь от второстепенных.[4]

В основе первого этапа статистического изучения связи лежит качественный анализ, связанный с анализом природы социального или экономического явления методами экономической теории, социологии, конкретной экономики. Второй этап - построение модели связи, базируется на методах статистики: группировках, средних величинах, и так далее. Третий, последний этап - интерпретация результатов, вновь связан с качественными особенностями изучаемого явления. Статистика разработала множество методов изучения связей. Выбор метода изучения связи зависит от познавательной цели и задач исследования.

Признаки по их сущности и значению для изучения взаимосвязи делятся на два класса. Признаки, обуславливающие изменения других, связанных с ними признаков, называются факторными, или просто факторами. Признаки, изменяющиеся под действием факторных признаков, называются результативными.[10]

В статистике различают функциональную и стохастическую зависимости. Функциональной называют такую связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака.

Если причинная зависимость проявляется не в каждом отдельном случае, а в общем, среднем при большом числе наблюдений, то такая зависимость называется стохастической. Частным случаем стохастической связи является корреляционная связь, при которой изменение среднего значения результативного признака обусловлено изменением факторных признаков.

Связи между явлениями и их признаками классифицируются по степени тесноты, направлению и аналитическому выражению.

По направлению выделяют связь прямую и обратную. Прямая - это связь, при которой с увеличением или с уменьшением значений факторного признака происходит увеличение или уменьшение значений результативного признака. Так, рост объемов производства способствует увеличению прибыли предприятия. В случае обратной связи значения результативного признака изменяются под воздействием факторного, но в противоположном направлении по сравнению с изменением факторного признака, то есть обратная - это связь, при которой с увеличением или с уменьшением значений одного признака происходит уменьшение или увеличение значений другого признака. Так снижение себестоимости единицы производимой продукции влечет за собой рост рентабельности.[1]

По аналитическому выражению выделяют связи прямолинейные (или просто линейные) и нелинейные. Если статистическая связь между явлениями может быть приблизительно выражена уравнением прямой линии, то ее называют линейной связью вида:

(1.1)



Если же связь может быть выражена уравнением какой-либо кривой линии, то такую связь называют нелинейной или криволинейной, например:

- параболы: (1.2)



- гиперболы: (1.3)



Для выявления наличия связи, ее характера и направления в статистике используются методы: приведения параллельных данных; графический; аналитических группировок; корреляции, регрессии.

Метод приведения параллельных данных основан на сопоставлении двух или нескольких рядов статистических величин. Такое сопоставление позволяет установить наличие связи и получить представление о ее характере.

Графически взаимосвязь двух признаков изображается с помощью поля корреляции. В системе координат на оси абсцисс откладываются значения факторного признака, а на оси ординат - результативного. Каждое пересечение линий, проводимых через эти оси, обозначаются точкой. При отсутствии тесных связей имеет место беспорядочное расположение точек на графике. Чем сильнее связь между признаками, тем теснее будут группироваться точки вокруг определенной линии, выражающей форму связи.



Рис.1.1. График корреляционного поля

В статистике принято различать следующие варианты зависимостей:

Парная корреляция - связь между двумя признаками (результативным и факторным, или двумя факторными).[7]

Частная корреляция - зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков.

Множественная корреляция - зависимость результативного и двух или более факторных признаков, включенных в исследование.

Корреляционный анализ имеет своей задачей количественное определение тесноты и направления связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи).

Теснота связи количественно выражается величиной коэффициентов корреляции, которые давая количественную характеристику тесноты связи между признаками, позволяют определять "полезность" факторных признаков при построении уравнения множественной регрессии. Знаки при коэффициентах корреляции характеризуют направление связи между признаками.

Регрессия тесно связана с корреляцией и позволяет исследовать аналитическое выражение взаимосвязи между признаками.

Регрессионный анализ заключается в определении аналитического выражения связи, в котором изменение одной величины (называемой зависимой или результативным признаком), обусловлено влиянием одной или нескольких независимых величин (факторных признаков).[3]

Одной из проблем построения уравнений регрессии является их размерность, то есть определение числа факторных признаков, включаемых в модель. Их число должно быть оптимальным. Сокращение размерности за счет исключения второстепенных, несущественных факторов позволяет получить модель, быстрее и качественнее реализуемую. В то же время, построение модели малой размерности может привести к тому, что она будет недостаточно полно описывать исследуемое явление или процесс.

При построении моделей регрессии должны соблюдаться следующие требования:

1. Совокупность исследуемых исходных данных должна быть однородной и математически описываться непрерывными функциями.
2. Возможность описания моделируемого явления одним или несколькими уравнениями причинно-следственных связей.
3. Все факторные признаки должны иметь количественное (числовое) выражение.
4. Наличие достаточно большого объема исследуемой выборочной совокупности.
5. Причинно-следственные связи между явлениями и процессами должны описываться линейной или приводимой к линейной форме зависимостью.
6. Отсутствие количественных ограничений на параметры модели связи.
7. Постоянство территориальной и временной структуры изучаемой совокупности.

Соблюдение данных требований позволяет построить модель, наилучшим образом описывающую реальные социально-экономические явления и процессы.[13]

2. ХАРАКТЕРИСТИКА РЕГРЕССИОННОГО АНАЛИЗА

2.1. Оценка взаимосвязи между факторным и результативным признаком на основе регрессионного анализа

Статистика разработала множество методов изучения связей, выбор которых зависит от целей исследования и от поставленных задач. Связи между признаками и явлениями, ввиду их большого разнообразия, классифицируются по ряду оснований. Признаки по их значению для изучения взаимосвязи делятся на 2 класса. Признаки, обуславливающие изменения других, связанных с ними признаков, называются факторными, или просто факторами. Признаки, изменяющиеся под действием факторных признаков, являются результативными. Связи между явлениями и их признаками классифицируются по степени тесноты связи, направлению и аналитическому выражению.[12]

Между различными явлениями и их признаками необходимо прежде всего выделить 2 типа связей: функциональную (жестко детерминированную) и статистическую (стохастически детерминированную).

В соответствии с жестко детерминистическим представлением о функционировании экономических систем необходимость и закономерность однозначно проявляются в каждом отдельном явлении, то есть любое действие вызывает строго определенный результат; случайными (непредвиденными заранее) воздействиями при этом пренебрегают. Поэтому при заданных начальных условиях состояние такой системы может быть определено с вероятностью, равной 1. Разновидностью такой закономерности является функциональная связь.

Связь признака у с признаком х называется функциональной, если каждому возможному значению независимого признака х соответствует 1 или несколько строго определенных значений зависимого признака у. Определение функциональной связи может быть легко обобщено для случая многих признаков х1,х2 …хn.[9]

Характерной особенностью функциональных связей является то, что в каждом отдельном случае известен полный перечень факторов, определяющих значение зависимого (результативного) признака, а также точный механизм их влияния, выраженный определенным уравнением.

В реальной общественной жизни ввиду неполноты информации жестко детерминированной системы, может возникнуть неопределенность, из-за которой эта система по своей природе должна рассматриваться как вероятностная, при этом связь между признаками становится стохастической.

Стохастическая связь – это связь между величинами, при которой одна из них, случайная величина у, реагирует на изменение другой величины х или других величин х1,х2 …хn(случайных или неслучайных) изменением закона распределения. Это обуславливается тем, что зависимая переменная (результативный признак), кроме рассматриваемых независимых, подвержена влиянию ряда неучтенных или неконтролируемых (случайных) факторов, а также некоторых неизбежных ошибок измерения переменных. Поскольку значения зависимой переменной подвержены случайному разбросу, они не могут быть предсказаны с достаточной точностью, а только указаны с определенной вероятностью.

Характерной особенностью стохастических связей является то, что они проявляются во всей совокупности, а не в каждой ее единице. Причём неизвестен ни полный перечень факторов, определяющих значение результативного признака, ни точный механизм их функционирования и взаимодействия с результативным признаком. Всегда имеет место влияние случайного. Появляющиеся различные значения зависимой переменной – реализация случайной величины.[5]

Проявление стохастических связей подвержено действию закона больших чисел: лишь в достаточно большом числе единиц индивидуальные особенности сгладятся, случайности взаимопогасятся, и зависимость, если она имеет существенную силу, проявится достаточно отчётливо.

Корреляционная связь существует там, где взаимосвязанные явления характеризуются только случайными величинами. При такой связи среднее значение (математическое ожидание) случайной величины результативного признака у закономерно изменяется в зависимости от изменения другой величины х или других случайных величин х1,х2 …хn. Корреляционная связь проявляется не в каждом отдельном случае, а во всей совокупности в целом. Только при достаточно большом количестве случаев каждому значению случайного признака х будет соответствовать распределение средних значений случайного признака у. Наличие корреляционных связей присуще многим общественным явлениям.

Корреляционная связь – понятие более узкое, чем стохастическая связь. Последняя может отражаться не только в изменении средней величины, но и в вариации одного признака в зависимости от другого, то есть любой другой характеристики вариации. Таким образом, корреляционная связь является частным случаем стохастической связи.[8]

Прямые и обратные связи. В зависимости от направления действия, функциональные и стохастические связи могут быть прямые и обратные. При прямой связи направление изменения результативного признака совпадает с направлением изменения признака-фактора, то есть с увеличением факторного признака увеличивается и результативный, и, наоборот, с уменьшением факторного признака уменьшается и результативный признак. В противном случае между рассматриваемыми величинами существуют обратные связи. Например, чем выше квалификация рабочего (разряд), тем выше уровень производительности труда – прямая связь. А чем выше производительность труда, тем ниже себестоимость единицы продукции – обратная связь.

Прямолинейные и криволинейные связи. По аналитическому выражению (форме) связи могут быть прямолинейными и криволинейными. При прямолинейной связи с возрастанием значения факторного признака происходит непрерывное возрастание (или убывание) значений результативного признака. Математически такая связь представляется уравнением прямой, а графически – прямой линией. Отсюда ее более короткое название – линейная связь. При криволинейных связях с возрастанием значения факторного признака возрастание (или убывание) результативного признака происходит неравномерно, или же направление его изменения меняется на обратное. Геометрически такие связи представляются кривыми линиями (гиперболой, параболой и т.д.).[14]

Однофакторные и многофакторные связи. По количеству факторов, действующих на результативный признак, связи различаются: однофакторные (один фактор) и многофакторные (два и более факторов). Однофакторные (простые) связи обычно называются парными (т.к. рассматривается пара признаков). Например, корреляционная связь между прибылью и производительностью труда. В случае многофакторной (множественной) связи имеют в виду, что все факторы действуют комплексно, то есть одновременно и во взаимосвязи. Например, корреляционная связь между производительностью труда и уровнем организации труда, автоматизации производства, квалификации рабочих, производственным стажем, простоями и другими факторными признаками. С помощью множественной корреляции можно охватить весь комплекс факторных признаков и объективно отразить существующие множественные связи.[12]

Для исследования стохастических связей широко используется метод сопоставления двух параллельных рядов, метод аналитических группировок, корреляционный анализ, регрессионный анализ и некоторые непараметрические методы.

Метод сопоставления двух параллельных рядов является одним из простейших методов. Для этого факторы, характеризующие результативный признак располагают в возрастающем или убывающем порядке (в зависимости от эволюции процесса и цели исследования), а затем прослеживают изменение величины результативного признака. Сопоставление и анализ расположенных таким образом рядов значений изучаемых величин позволяют установить наличие связи и ее направление. Зависимость между факторами и показателями может прослеживаться во времени (параллельные динамические ряды).

Метод аналитических группировок тоже относится к простейшим методам. Чтобы выявить зависимость с помощью этого метода, нужно произвести группировку единиц совокупности по факторному признаку и для каждой группы вычислить среднее или относительное значение результативного признака. Сопоставляя затем изменения результативного признака по мере изменения факторного можно выявить направление, характер и тесноту связи между ними.[14]

В общем виде задача статистики в области изучения взаимосвязей состоит не только в количественной оценке их наличия, направления и силы связи, но и в определении формы (аналитического выражения) влияния факторных признаков на результативный. Для ее решения применяют методы корреляционного и регрессионного анализа.

Задачи корреляционного анализа сводятся к измерению тесноты известной связи между варьирующими признаками, определению неизвестных причинных связей (причинный характер которых должен быть выяснен с помощью теоретического анализа) и оценки факторов, оказывающих наибольшее влияние на результативный признак.

Задачами регрессионного анализа являются выбор типа модели (формы связи), установление степени влияния независимых переменных на зависимую и определение расчётных значений зависимой переменной (функции регрессии).

Решение всех названных задач приводит к необходимости комплексного использования этих методов.

Корреляционный и регрессионный анализ. Исследование связей в условиях массового наблюдения и действия случайных факторов осуществляется, как правило, с помощью экономико-статистических моделей. В широком смысле модель – это аналог, условный образ (изображение, описание, схема, чертёж и т.п.) какого-либо объекта, процесса или события, приближенно воссоздающий «оригинал». Модель представляет собой логическое или математическое описание компонентов и функций, отображающих существенные свойства моделируемого объекта или процесса, даёт возможность установить основные закономерности изменения оригинала. В модели оперируют показателями, исчисленными для качественно однородных массовых явлений (совокупностей). Выражение и модели в виде функциональных уравнений используют для расчёта средних значений моделируемого показателя по набору заданных величин и для выявления степени влияния на него отдельных факторов.[12]

По количеству включаемых факторов модели могут быть однофакторными и многофакторными (два и более факторов).

В зависимости от познавательной цели статистические модели подразделяются на структурные, динамические и модели связи.

Двухмерная линейная модель корреляционного и регрессионного анализа (однофакторный линейный корреляционный и регрессионный анализ). Наиболее разработанной в теории статистики является методология так называемой парной корреляции, рассматривающая влияние вариации факторного анализа х на результативный признак у и представляющая собой однофакторный корреляционный и регрессионный анализ. Овладение теорией и практикой построения и анализа двухмерной модели корреляционного и регрессионного анализа представляет собой исходную основу для изучения многофакторных стохастических связей.[3]

Важнейшим этапом построения регрессионной модели (уравнения регрессии) является установление в анализе исходной информации математической функции. Сложность заключается в том, что из множества функций необходимо найти такую, которая лучше других выражает реально существующие связи между анализируемыми признаками. Выбор типов функции может опираться на теоретические знания об изучаемом явлении, опят предыдущих аналогичных исследований, или осуществляться эмпирически – перебором и оценкой функций разных типов и т.п.

При изучении связи экономических показателей производства (деятельности) используют различного вида уравнения прямолинейной и криволинейной связи. Внимание к линейным связям объясняется ограниченной вариацией переменных и тем, что в большинстве случаев нелинейные формы связи для выполнения расчётов преобразуют (путём логарифмирования или замены переменных) в линейную форму. Уравнение однофакторной (парной) линейной корреляционной связи имеет вид:

ŷ = a0 + a1x ,

где ŷ - теоретические значения результативного признака, полученные по уравнению регрессии;

a0, a1 - коэффициенты (параметры) уравнения регрессии.

Поскольку a0 является средним значением у в точке х=0, экономическая интерпретация часто затруднена или вообще невозможна.

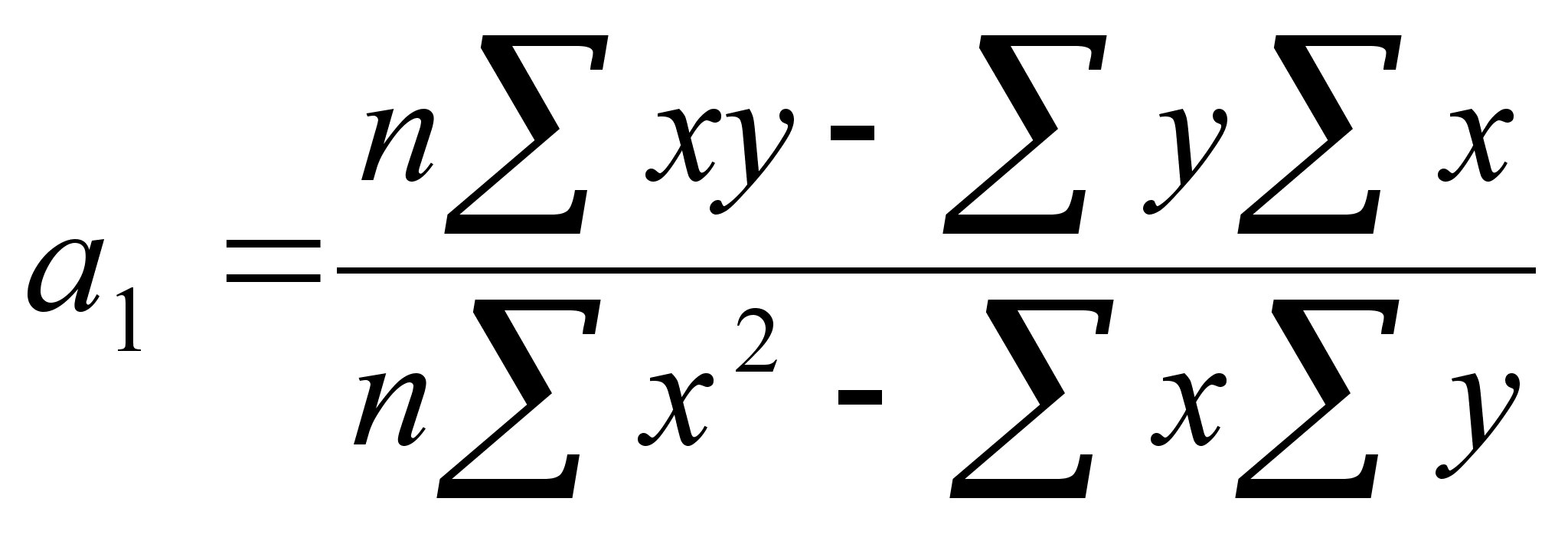
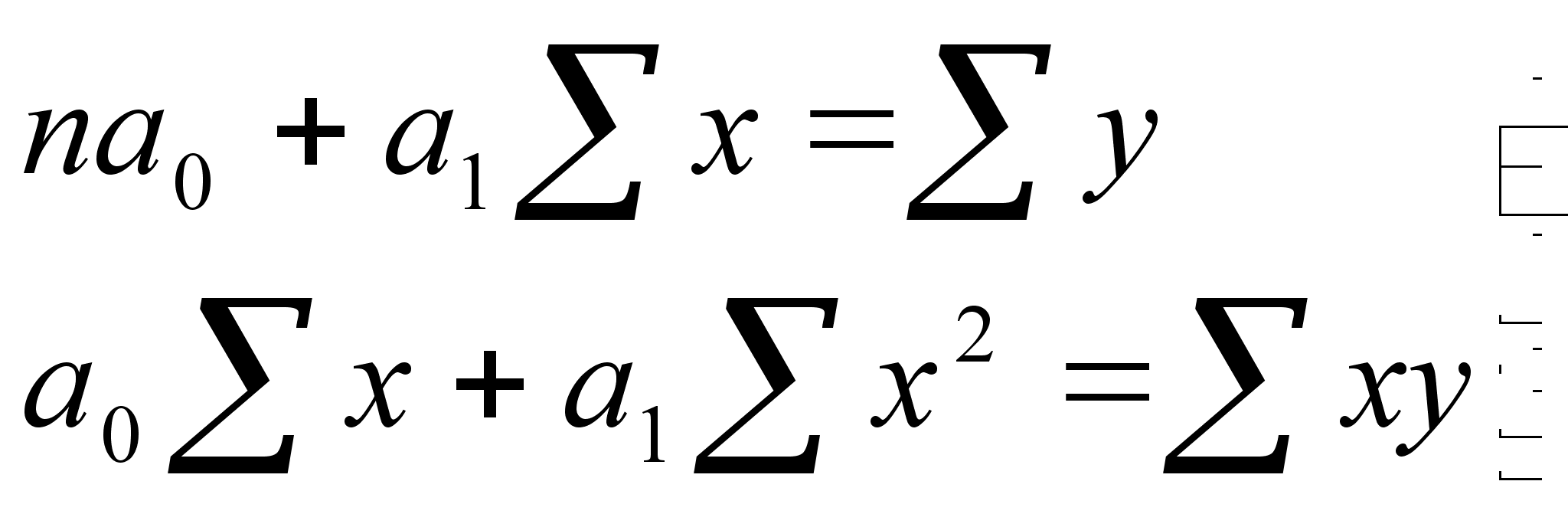
Коэффициент парной линейной регрессии a1имеет смысл показателя силы связи между вариацией факторного признака х и вариацией результативного признака у. Вышеприведенное уравнение показывает среднее значение изменения результативного признака у при изменении факторного признака х на одну единицу его измерения, то есть вариацию у, приходящуюся на единицу вариации х. Знак a1 указывает направление этого изменения.[7]

Параметры уравнения a0, a1 находят методом наименьших квадратов (метод решения систем уравнений, при котором в качестве решения принимается точка минимума суммы квадратов отклонений), то есть в основу этого метода положено требование минимальности сумм квадратов отклонений эмпирических данных yi от выровненных ŷ :

Σ(yi – ŷ)2 = Σ(yi – a0 – a1xi)2 → min

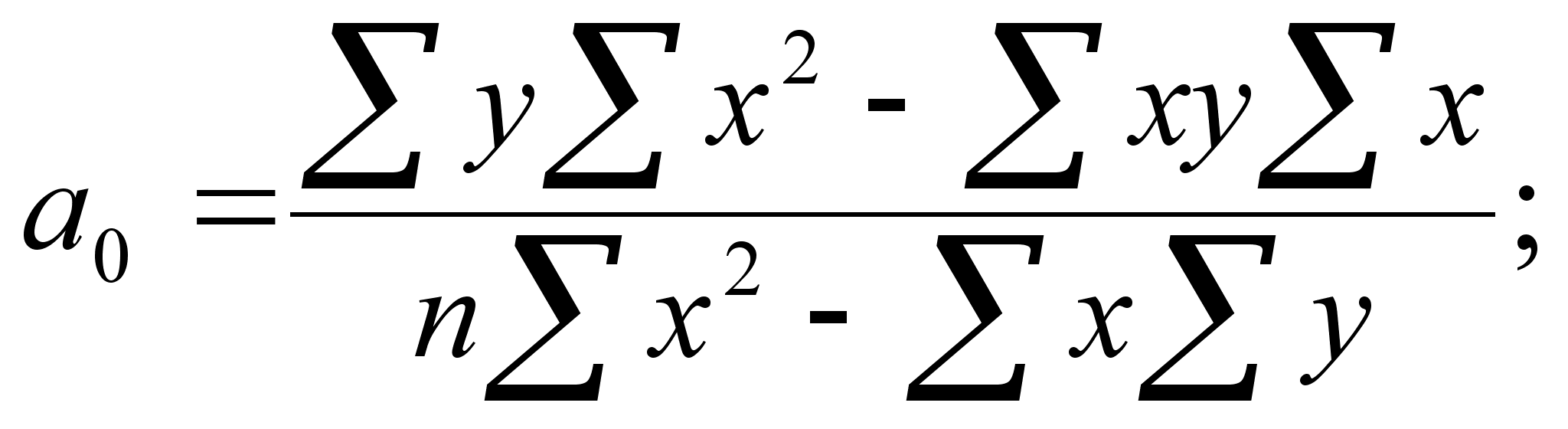
Для нахождения минимума данной функции приравняем к нулю ее частные производные и получим систему двух линейных уравнений, которая называется системой нормальных уравнений:

(2.1)

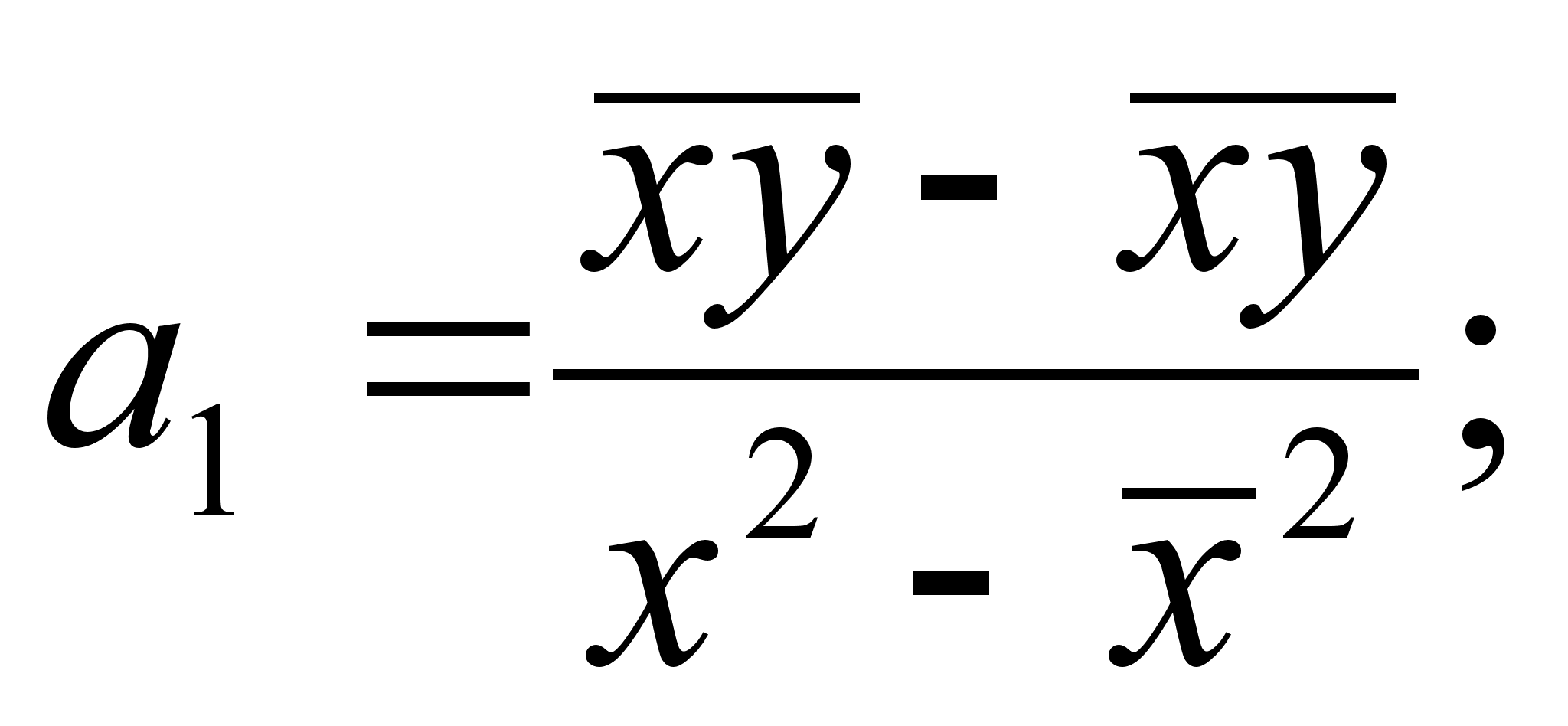
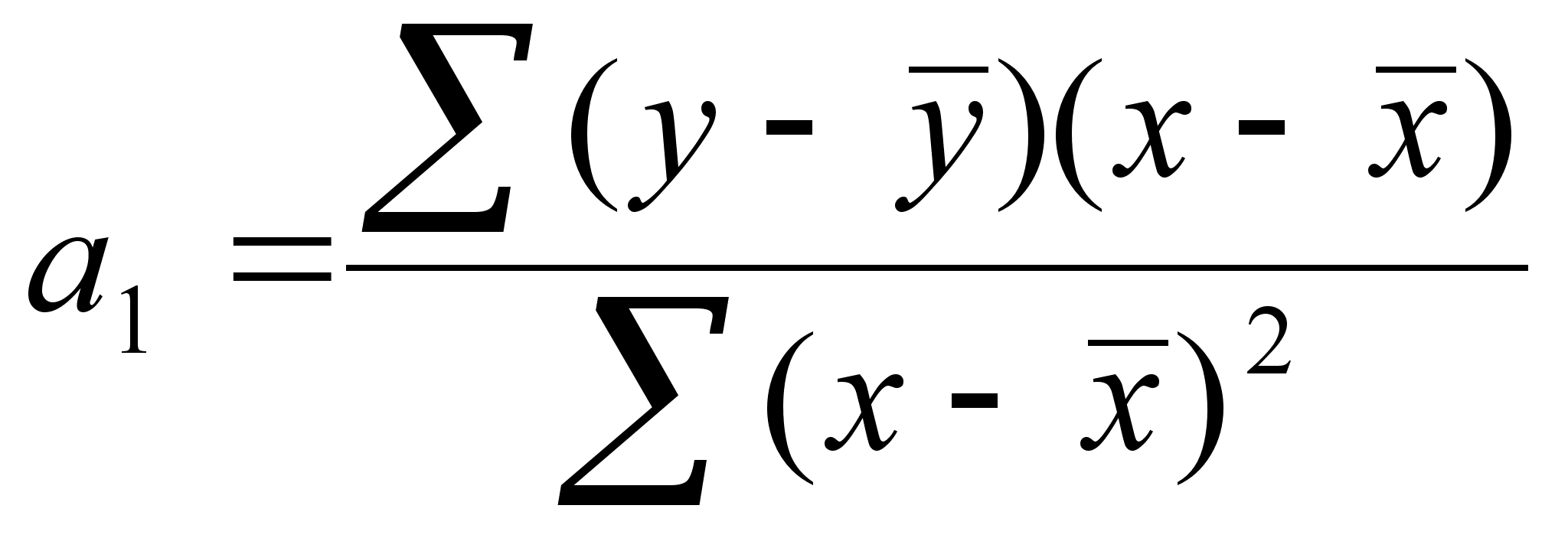


Решим эту систему уравнений в общем виде:

(2.2)



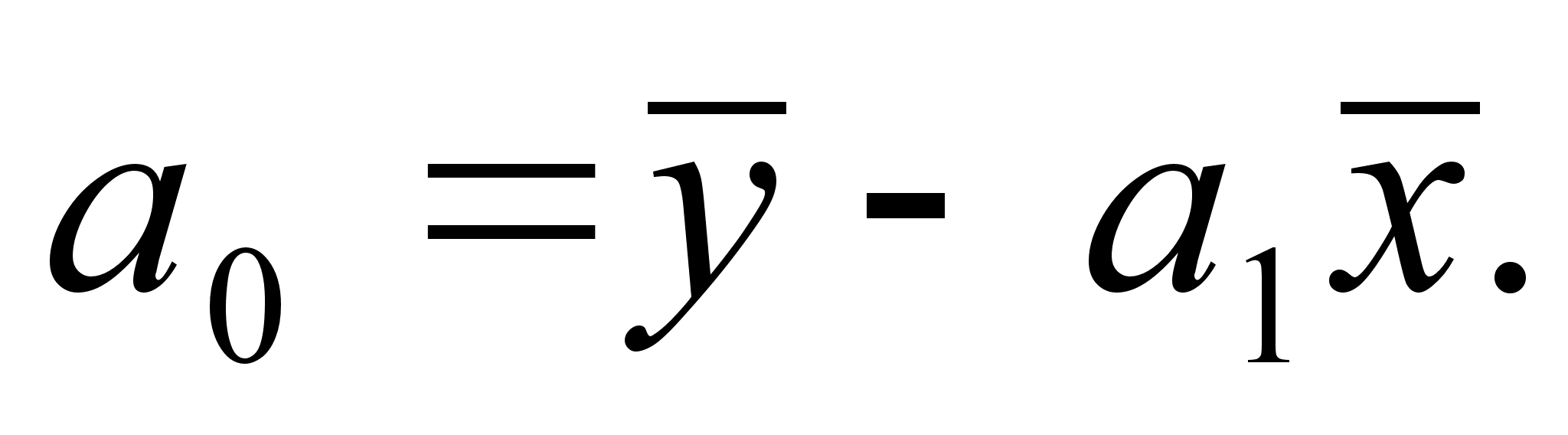
или



(2.3)

Параметры уравнения парной линейной регрессии иногда удобно исчислять по следующей формуле, дающим тот же результат:

(2.4)



Определив значения a0, a1и подставив их в уравнение связи ŷ = a0 + a1x , находим значения ŷ , зависящие только от заданного значения х.[11]

2.2. Отбор факторных признаков для построения множественной регрессионной модели

Регрессионный анализ наиболее широко используемый метод многомерного статистического анализа. Термин ''множественная регрессия'' объясняется тем, что анализу подвергается зависимость одного признака (результирующего) от набора независимых (факторных) признаков. Разделение признаков на результирующий и факторные осуществляется исследователем на основе содержательных представлений об изучаемом явлении (процессе). Все признаки должны быть количественными (хотя допускается и использование дихотомических признаков, принимающих лишь два значения, например 0 и 1).

Для корректного использования регрессионного анализа требуется выполнение определенных условий. Факторные признаки должны быть некоррелированы (отсутствие мультиколлинеарности), они предполагаются замеренными точно и в их измерениях нет автокорреляции, т.е. значения признаков у одного объекта не должны зависеть от значений признаков у других объектов.[5]

Изучение связи между результативным и двумя или более факторными признаками называется множественной регрессией. При исследовании зависимостей методами множественной регрессии ставят 2 задачи.

1. определение аналитического выражения связи между результативным признаком у и фактическими признаками х1, х2, х3, …хк, т.е. найти функцию у=f(х1, х2, …хк)
2. Оценка тесноты связи между результативным и каждым из факторных признаков.

Корреляционно-регрессионная модель (КРМ) – такое уравнение регрессии, которое включает основные факторы, влияющие на вариацию результативного признака.

#### Построение модели множественной регрессии включает этапы:

1. выбор формы связи
2. отбор факторных признаков
3. обеспечение достаточного объема совокупности для получения верных оценок.
4. все множество связей между переменными, встречающиеся на практике достаточно полно описывается функциями 5-ти видов:
5. линейная: 
6. степенная: 
7. показательная: 
8. парабола: 
9. гипербола: 

Хотя все 5 функций присутствуют в практике КРА, наиболее часто используется линейная зависимость, как наиболее простая и легко поддающаяся интерпретации уравнение линейной зависимости: , (2.5)

где к – множество факторов включающихся в уравнение,

bj – коэффициент условно-чистой регрессии, который показывает среднее по совокупности отклонение результативного признака от его среднего значения при отклонении фактора xj от своей средней величины на единицу при условии, что все остальные факторы, входящие в уравнение сохраняют средние значения.[9]

Параметры уравнения множественной регрессии и определение с помощью МНК.



Пример:

Потребление йогурта

Доход

Возраст

Пол

Социальное положение

0,72

0,64

0,4

0,93

0

0

0 – т.к. >0,7 следовательно на них обращаем особое внимание ЭКО. Шкала тесноты связи:

Если связь 0 – 0,3 – слабая связь

0,3 – 0,5 – заметная

0,3 – 0,5 – тесная

0,7 – 0,9 – высокая

более 0,9 – весьма высокая

Затем сравниваем два признака (доход и пол) <0,7, то включаем в уравнение множественной регрессии.

Отбор факторов для включения в уравнение множественной регрессии:

1. между результативным и фактическим признаками должна быть причинно-следственная зависимость.
2. результативный и фактический признаки должны быть тесно связаны между собой иначе возникает явление мультиколлинеарности (>06), т.е. включенные в уравнение факторные признаки влияют не только на результативный, но друг на друга, что влечет к неверной интерпретации числовых данных.

## Отбор факторов для модели осуществляется в два этапа. На первом идет анализ, по результатам которого исследователь делает вывод о необходимости рассмотрения тех или иных явлений в качестве переменных, определяющих закономерности развития исследуемого процесса, на втором - состав предварительно отобранных факторов уточняется непосредственно по результатам статистического анализа.[1]

Методы отбора факторов для включения в уравнение множественной регрессии:

1. экспертный метод – основан на интуитивно логическом анализе который выполняется высококвалифицированными экспертами.
2. использование матриц парных коэффициентов корреляции осуществляется параллельно с первым методом, матрица симметрична относительно единичной диагонали.
3. пошаговый регрессионный анализ – последовательное включение факторных признаков в уравнение регрессии и проверки значимости проводится на основании значений двух показателей на каждом шаге. Показатель корреляции, регрессии.

Показатель корреляции: рассчитывают изменение теоретической корреляции отношения или изменение средней остаточной дисперсии. Показатель регрессии – изменение коэффициента условно чистой регрессии.[14]

2.3. Проверка адекватности моделей, построенных на основе уравнений регрессии

Для практического использования моделей регрессии большое значение имеет их адекватность, т.е. соответствие фактическим статистическим данным.

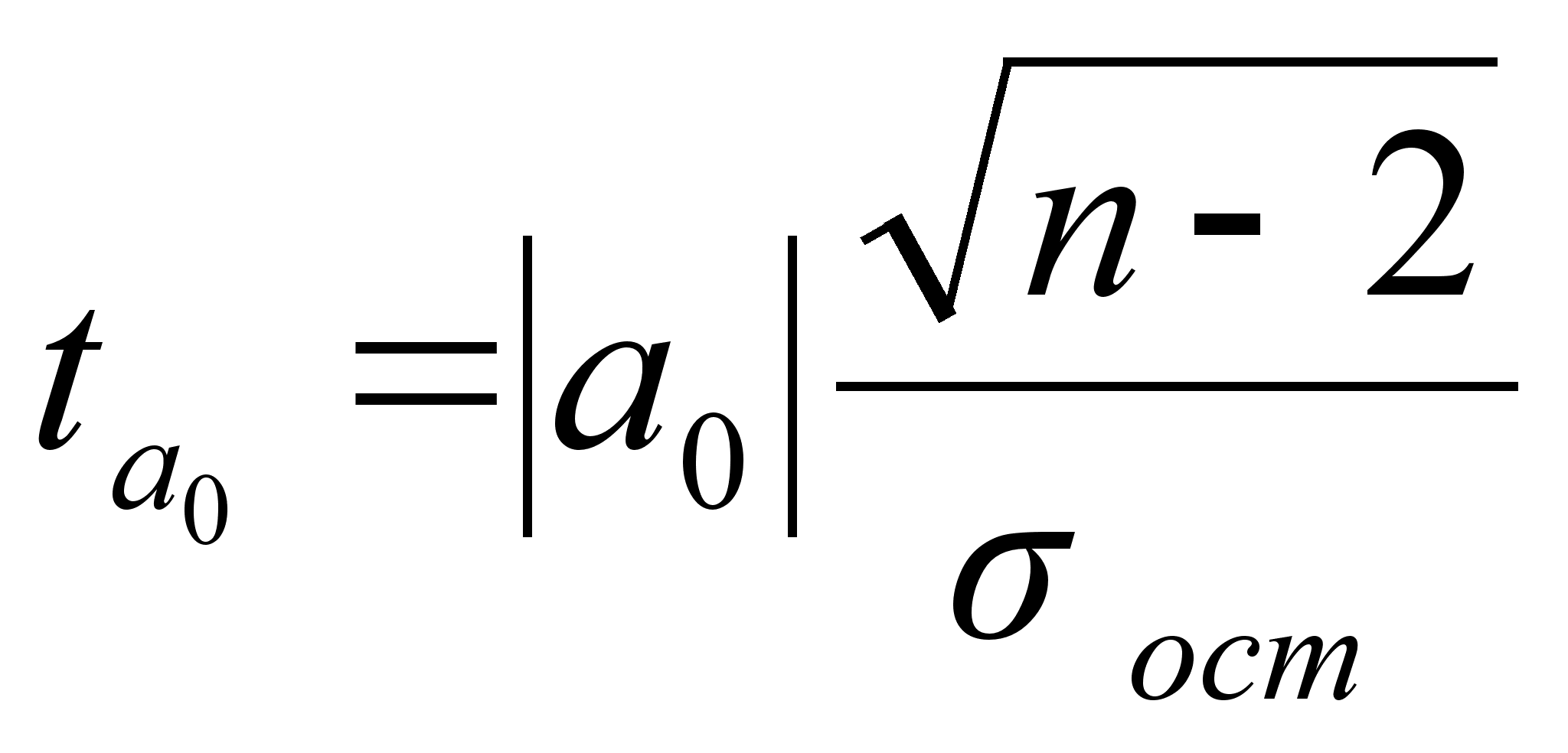
Корреляционный и регрессионный анализ обычно (особенно в условиях так называемого малого и среднего бизнеса) проводится для ограниченной по объёму совокупности. Поэтому показатели регрессии и корреляции – параметры уравнения регрессии, коэффициенты корреляции и детерминации могут быть искажены действием случайных факторов. Чтобы проверить, насколько эти показатели характерны для всей генеральной совокупности, не являются ли они результатом стечения случайных обстоятельств, необходимо проверить адекватность построенных статистических моделей.

Проверка адекватности моделей, построенных на основе уравнений регрессии, начинается с проверки значимости каждого коэффициента регрессии.[12]

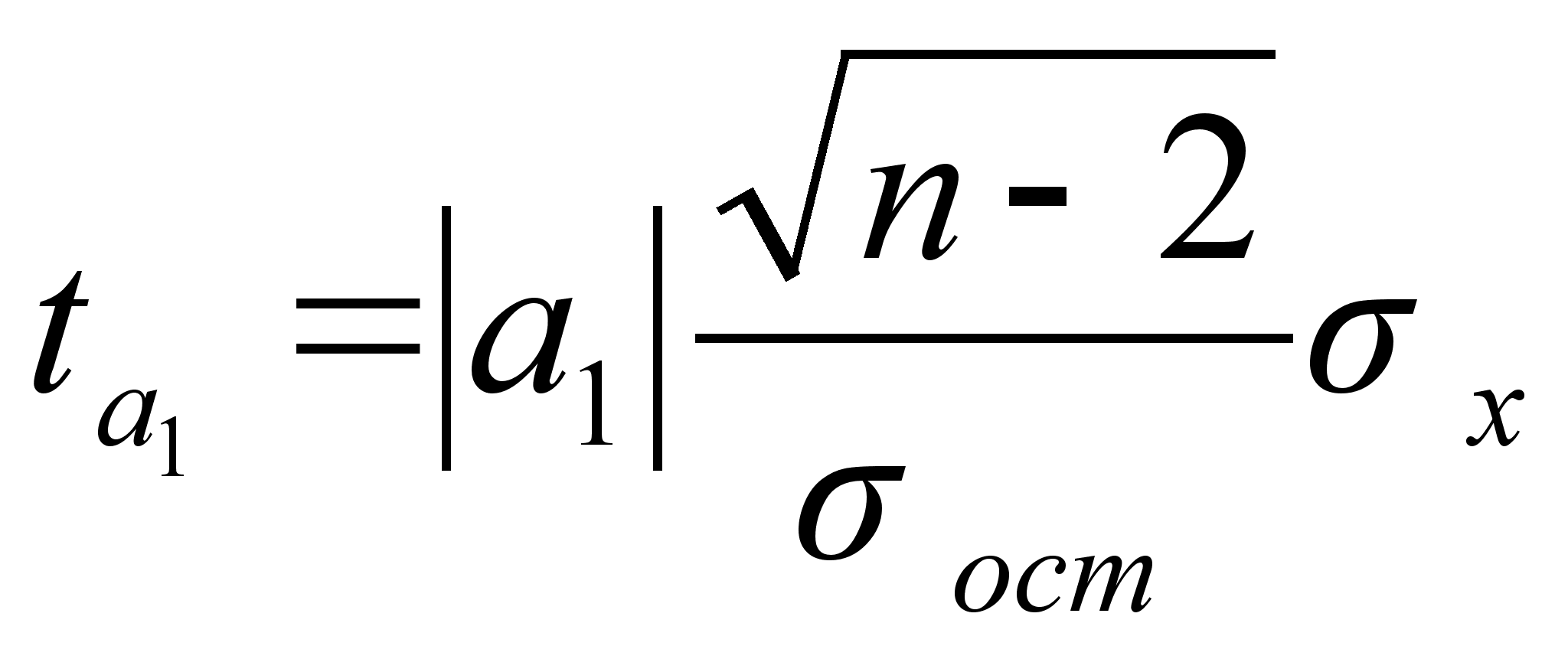
При численности объектов анализа до 30 единиц возникает необходимость проверки значимости (существенности) каждого коэффициента регрессии. При этом выясняют насколько вычисленные параметры характерны для отображения комплекса условий: не являются ли полученные значения параметров результатами действия случайных причин.

Значимость коэффициентов простой линейной регрессии (применительно к совокупностям, у которых n<30) осуществляют с помощью t-критерия Стьюдента. При этом вычисляют расчетные (фактические) значения t-критерия.

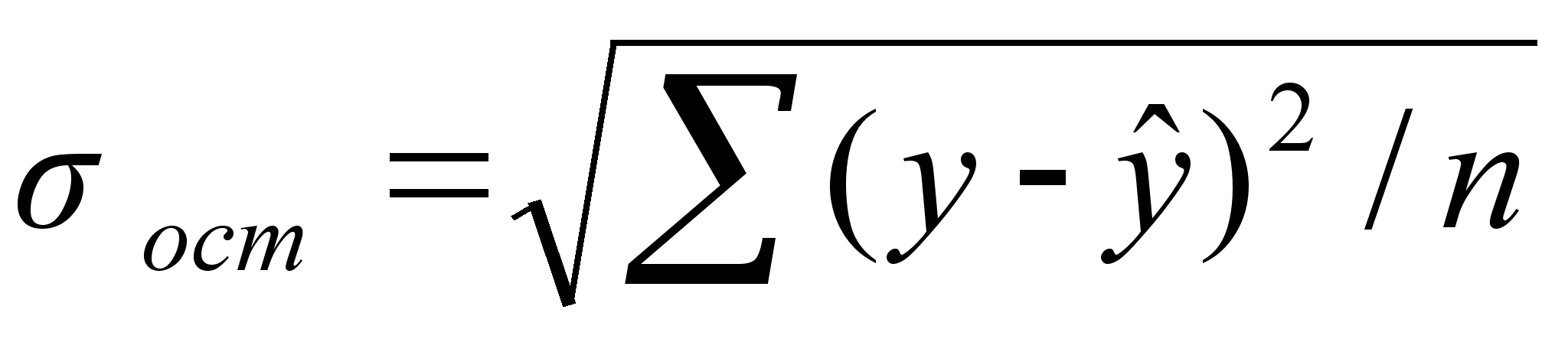
(2.6)  
Для параметра **a0**:



(2.7)  
для параметра **a1**:

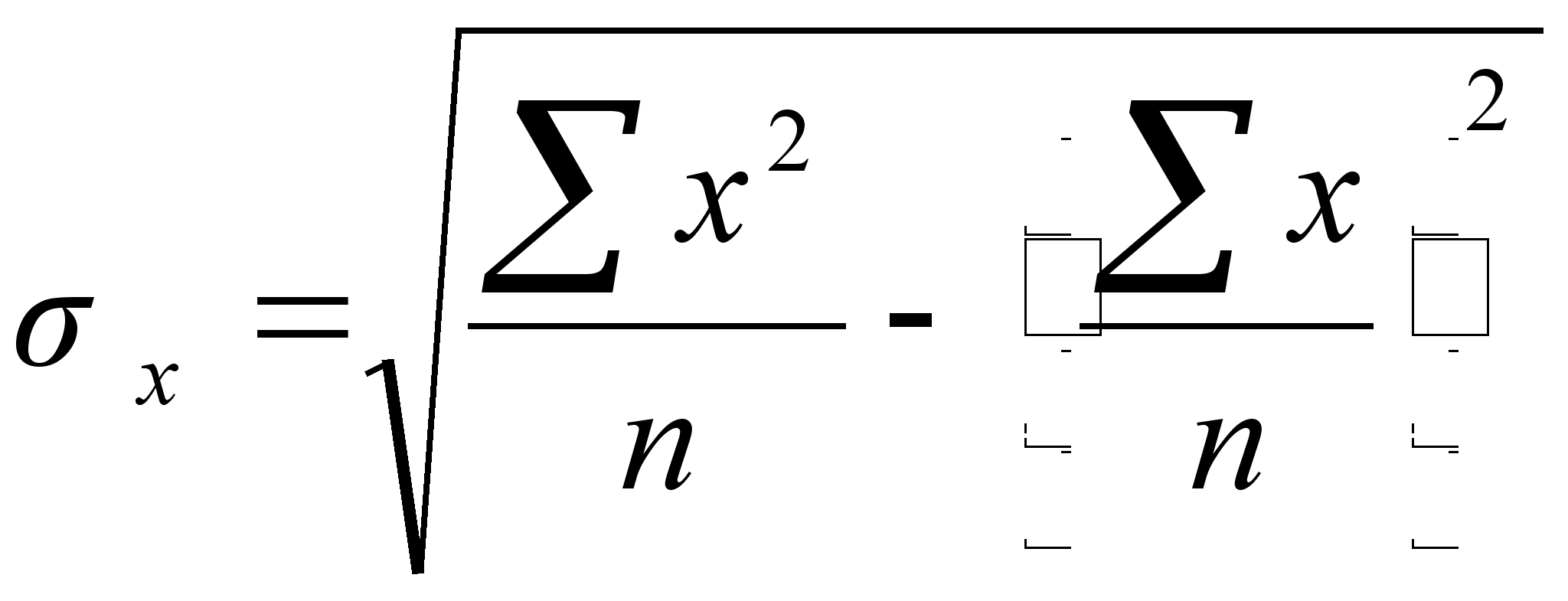
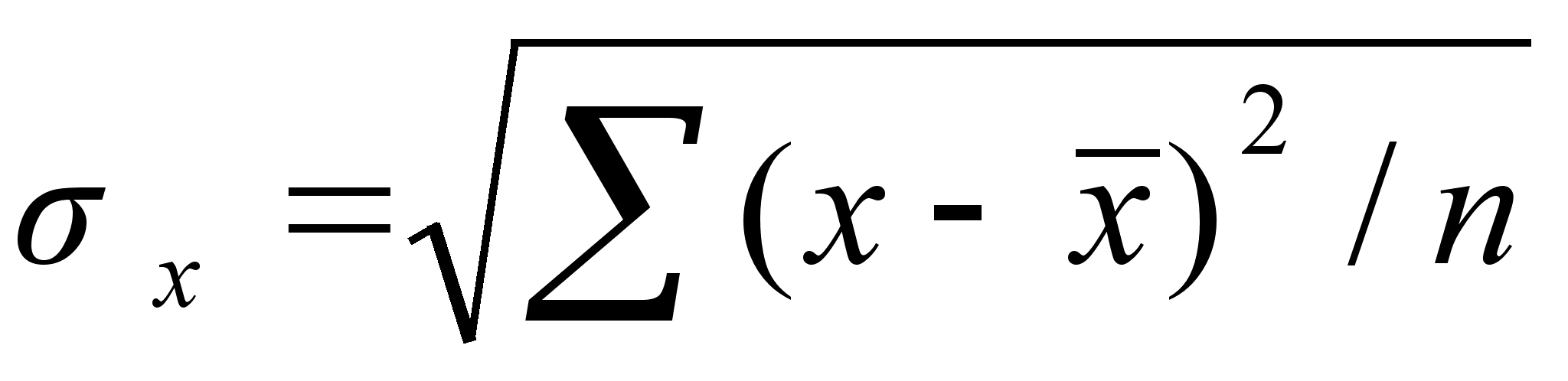


(2.8)  
где **n** - объём выборки;

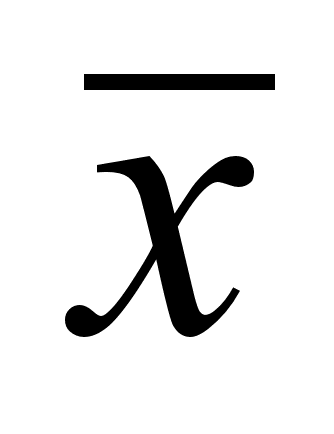


- среднее квадратическое отклонение результативного признака от выравненных значений **ŷ**;

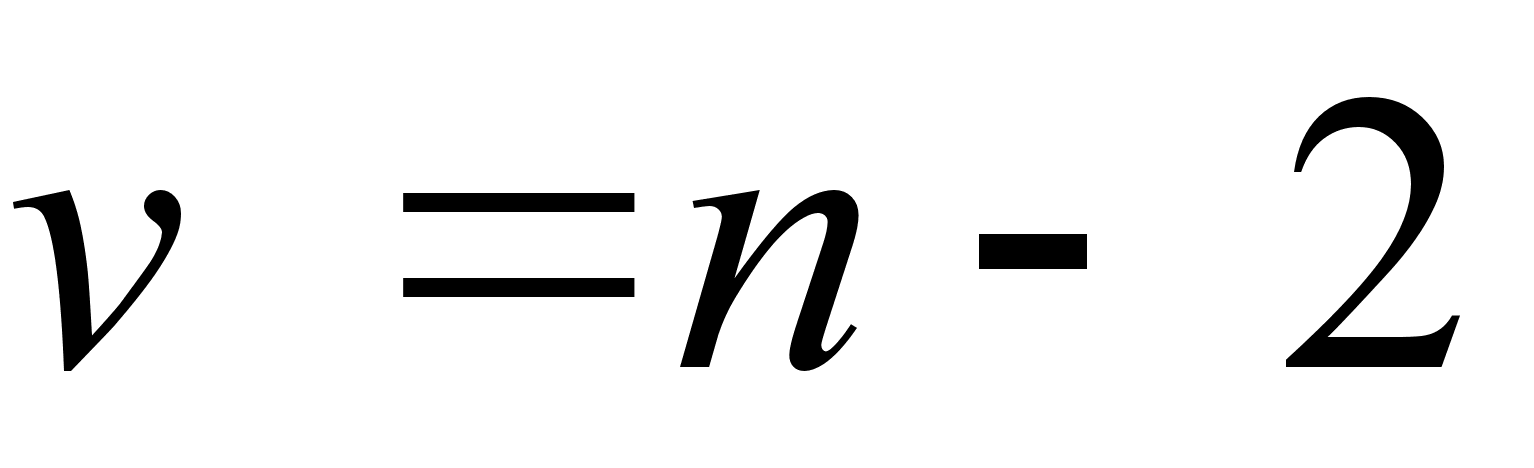
 или  (2.9)



- среднее квадратическое отклонение факторного признака **x** от общей средней .[8]



Вычисленные по вышеприведенным формулам значения сравнивают с критическими ***t***, которые определяют по таблице Стьюдента с учетом принятого уровня значимости *α*и числом степеней свободы вариации . В социально-экономических исследованиях уровень значимости α обычно принимают равным 0,05. Параметр признаётся значимым (существенным) при условии, если *tрасч*>*tтабл*. В таком случае практически невероятно, что найденные значения параметров обусловлены только случайными совпадениями.



Параметр модели признается статистически значимым, если tp>tкр

Наиболее сложным в этом выражении является определение дисперсии, которая может быть рассчитана двояким способом.

Наиболее сложным этапом, завершающим регрессионный анализ, является интерпретация уравнения, т.е. перевод его с языка статистики и математики на язык экономиста.

Интерпретация моделей регрессии осуществляется методами той отрасли знаний, к которой относятся исследуемые явления. Но всякая интерпретация начинается со статистической оценки уравнения регрессии в целом и оценки значимости входящих в модель факторных признаков, т.е. с выяснения, как они влияют на величину результативного признака. Чем больше величина коэффициента регрессии, тем значительнее влияние данного признака на моделируемый. Особое значение при этом имеет знак перед коэффициентом регрессии. Знаки коэффициентов регрессии говорят о характере влияния на результативный признак. Если факторный признак имеет знак плюс, то с увеличением данного фактора результативный признак возрастает; если факторный признак со знаком минус, то с его увеличением результативный признак уменьшается.[12]

Интерпретация этих знаков полностью определяется социально-экономическим содержанием моделируемого (результативного) признака. Если его величина изменяется в сторону увеличения, то плюсовые знаки факторных признаков имеют положительное влияние. При изменении результативного призна-л-1 в сторону снижения положительное значение имеют минусовые знаки факторных признаков. Если экономическая теория подсказывает, что факторный признак должен иметь положительное значение, а он со знаком минус, то необходимо проверить расчеты параметров уравнения регрессии. Такое явление чаще всего бывает в силу допущенных ошибок при решении. Однако следует иметь в виду, что при анализе совокупного влияния факторов, при наличии взаимосвязей между ними характер их влияния может меняться. Для того чтобы быть уверенным, что факторный признак изменил знак влияния, необходима тщательная проверка решения данной модели, так как часто знаки могут меняться в силу допустимых ошибок при сборе или обработке информации.[4]

При адекватности уравнения регрессии исследуемому процессу возможны следующие варианты.

1. Построенная модель на основе ее проверки по F-критерию Фишера в целом адекватна, и все коэффициенты регрессии значимы. Такая модель может быть использована для принятия решений к осуществлению прогнозов.

2. Модель по F-критерию Фишера адекватна, но часть коэффициентов регрессии незначима. В этом случае модель пригодна для принятия некоторых решений, но не для производства прогнозов.

3. Модель по F-критерию Фишера адекватна, но все коэффициенты регрессии незначимы. Поэтому модель полностью считается неадекватной. На ее основе не принимаются решения и не осуществляются прогнозы.[12]

3. ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ ИЗУЧЕНИЯ ОБЪЕКТА ИССЛЕДОВАНИЯ

Таблица 3.1.

Исходные данные[15]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № предприятия | Объем добычи, тыс. т | Численность рабочих, чел | Стоимость основных фондов, тыс. у.е. | Фондоотдача, тыс. т/ | Производительность труда |
| 1 | 1275 | 240 | 10450 | 0,12 | 5,31 |
| 2 | 1250 | 210 | 7458 | 0,17 | 5,95 |
| 3 | 1054 | 264 | 9845 | 0,11 | 4 |
| 4 | 1513 | 240 | 8580 | 0,18 | 6,3 |
| 5 | 2214 | 276 | 9900 | 0,22 | 8,02 |
| 6 | 950 | 234 | 9790 | 0,1 | 4,06 |
| 7 | 1890 | 246 | 9350 | 0,2 | 7,68 |
| 8 | 2380 | 250 | 12650 | 0,19 | 9,52 |
| 9 | 2065 | 300 | 11290,4 | 0,18 | 6,88 |
| 10 | 1785 | 276 | 13200 | 0,14 | 6,47 |
| 11 | 1420 | 290 | 11200 | 0,13 | 4,9 |
| 12 | 1720 | 310 | 1460 | 1,18 | 5,55 |

Выделим основные показатели деятельности предприятия такие как: объем добычи, стоимость ОПФ и производительность труда рабочих.

Для начала необходимо найти недостающие данные. Проведём расчет производительности труда по формуле 3.1:

, (3.1)



где - выпуск валовой продукции;



- средняя списочная численность.

Результаты расчетов поместим в таблицу 3.2.

Таблица 3.2

Производительность труда, млн. руб./чел.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Пред-ие | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Производ-ность  труда | 5,31 | 5,95 | 4 | 6,3 | 8,02 | 4,06 | 7,68 | 9,52 | 6,88 | 6,47 | 4,9 | 5,55 |

Для установления зависимости объёма добычи от стоимости ОПФ и производительности труда рабочих проведём корреляционно-регрессионный анализ.

Чтобы провести анализ, присвоим переменным C, X, Y значения показателей величин: стоимости ОПФ, производительность труда рабочих и объёма добычи соответственно.

Прежде всего найдём средние значения показателей объёма добычи, стоимости основных фондов и производительности труда (X, Y, Z) по формуле средней арифметической 3.2:

, (3.2)

где xi –значение показателя i-го предприятия;

1. количество предприятий.

Результат представим в виде таблицы 3.3

Таблица 3.3

Средние значения показателей

|  |  |  |  |
| --- | --- | --- | --- |
|  | Y  (oбъем добычи), тыс. т. | X  (стоимость основных фондов), млн.руб. | Z (производительность труда), тыс. т./чел. |
| Cреднее значение показателей | 1626,33 | 9597,78 | 6,22 |

Далее проведём ряд расчётов и для удобства их проведения поместим результаты промежуточных расчетов в таблицу 3.4.

Таблица 3.4

Результаты расчетов

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Предприятие | Xi-X | (Xi-X)2 | Yi-Y | (Yi-Y)2 | Zi-Z | (Zi-Z)2 |
| 1 | 852,22 | 726273,25 | -351,33 | 123435,11 | -0,91 | 0,83 |
| 2 | 7458,00 | 55621764,00 | 1250,00 | 1562500,00 | 5,95 | 35,40 |
| 3 | 9845,00 | 96924025,00 | 1054,00 | 1110916,00 | 4,00 | 16,00 |
| 4 | 8580,00 | 73616400,00 | 1513,00 | 2289169,00 | 6,30 | 39,69 |
| 5 | 9900,00 | 98010000,00 | 2214,00 | 4901796,00 | 8,02 | 64,32 |
| 6 | 9790,00 | 95844100,00 | 950,00 | 902500,00 | 4,06 | 16,48 |
| 7 | 9350,00 | 87422500,00 | 1890,00 | 3572100,00 | 7,68 | 58,98 |
| 8 | 12650,00 | 160022500,00 | 2380,00 | 5664400,00 | 9,52 | 90,63 |
| 9 | 11290,40 | 127473132,16 | 2065,00 | 4264225,00 | 6,88 | 47,33 |
| 10 | 13200,00 | 174240000,00 | 1785,00 | 3186225,00 | 6,47 | 41,86 |
| 11 | 11200,00 | 125440000,00 | 1420,00 | 2016400,00 | 4,90 | 24,01 |
| 12 | 1460,00 | 2131600,00 | 1720,00 | 2958400,00 | 5,55 | 30,80 |
| Сумма |  | 1097472294,41 |  | 32552066,11 |  | 466,35 |

Найдём среднее квадратическое отклонение каждого показателя по формуле 3.3:

   , (3.3)



где – значение показателя х i-го предприятия;



- среднее значение показателя х i-го предприятия;



n - число предприятий

Среднее квадратическое отклонение, показывает, на сколько в среднем отклоняются конкретные варианты признака от среднего значения. Используя формулу 3.3, рассчитаем квадратические отклонения каждого показателя:

σ(x) = 9563,26

σ(y) = 1647,02

σ(z) = 6,23

Найдём парные коэффициенты корреляции по формуле 3.4:

, (3.4)

где -ковариация показателей х, y,

- среднее квадратическое отклонение показателя х;

- среднее квадратическое отклонение показателя у.

Корреляция — статистическая взаимосвязь двух или нескольких случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми). При этом, изменения одной или нескольких из этих величин приводят к систематическому изменению другой или других величин. Математической мерой корреляции двух случайных величин служит коэффициент корреляции.[12]

Рассчитаем ковариацию между парами показателей (Y,X), (Y,Z), (X,Z) по формуле 3.5:

, (3.5)

где  *Хi* -значение показателя х на i-м предприятие;

- среднее значение показателя Х;

*Yi* - значение показателя Y на i предприятие;

- cреднее значение показателя Y;

1. количество предприятий.

После того как все необходимые предварительные расчёты были произведены рассчитаем ковариацию и занесем данные в таблицу 3.5.

Таблица 3.5

Парные показатели ковариации

|  |  |  |  |
| --- | --- | --- | --- |
| Предприятие | (Xi-X)(Y1-Y) | (Zi-Z)(Yi-Y) | (Xi-X)(Zi-Z) |
| 1 | -299412,12 | 319,7133 | -775,5172 |
| 2 | 9322500 | 7437,5 | 44375,1 |
| 3 | 10376630 | 4216 | 39380 |
| 4 | 12981540 | 9531,9 | 54054 |
| 5 | 21918600 | 17756,28 | 79398 |
| 6 | 9300500 | 3857 | 39747,4 |
| 7 | 17671500 | 14515,2 | 71808 |
| 8 | 30107000 | 22657,6 | 120428 |
| 9 | 23314676 | 14207,2 | 77677,95 |
| 10 | 23562000 | 11548,95 | 85404 |
| 11 | 15904000 | 6958 | 54880 |
| 12 | 2511200 | 9546 | 8103 |
| Сov | 176670734 | 122551,3 | 674479,93 |

На основе полученных данных рассчитаем парные коэффициенты корреляции по формуле:

, (3.6)

где -ковариация показателей х, y;

- среднее квадратическое отклонение показателя х;

- среднее квадратическое отклонение показателя у.

r(y,x)= 0,93

r(y,z)= 0,99

r(x,z)= 0,94

Полученные коэффициенты корреляции указывают на весьма сильную связь каждого фактора с результатом, а также высокую межфакторную зависимость (факторы X и Z явно коллинеарны, т.к. r(x,z)= 0,94>0,7). При такой сильной межфакторной зависимости искажаются результаты, т.к. коэффициенты парной корреляции дают завышенные оценки тесноты связи. Именно по этой причине рекомендуется при наличии сильной коллинеарности (взаимосвязи) факторов исключать из исследования тот фактор, у которого теснота парной зависимости меньше, чем теснота межфакторной связи. В нашем случае это стоимость основных фондов (Х).

Для нахождения параметров линейного уравнения множественной регрессии  воспользуемся формулами:

; (3.7)

; (3.8)

 , (3.9)

где a = - 5,3825

b1= - 0,0047

b2 = 269,6281

Для практического использования моделей регрессии большое значение имеет их адекватность, т.е. соответствие фактическим статистическим данным.

Проведём проверку парных коэффициентов корреляции (X,Y), (Z,Y), (Z,X). При этом вычисляют расчетные (фактические) значения t-критерия Стьюдента проводится путем сопоставления их значений с величиной случайной ошибки:



, (3.10)

где r - парный коэффициент корреляции;

Mr - случайные ошибки коэффициента корреляции

Проведём расчёт квадрата коэффициента корреляции по формуле:



(3.11)

где b – коэффициент регрессии;

- значение показателя х на i-м предприятие;



- среднее значение показателя x;



- значение показателя y на i-м предприятии;



- cреднее значение показателя y.



Для этого найдём коэффициент регрессии по формуле:



, (3.12)

где b – коэффициент регрессии;

- значение показателя х на i-м предприятие;



- среднее значение показателя x;



- значение показателя y на i-м предприятии;



- cреднее значение показателя y.



Расчёты представим в таблице 3.6

Таблица 3.6

Квадраты коэффициентов регрессии и парных коэффициентов корреляции

|  |  |  |  |
| --- | --- | --- | --- |
| Пары показателей | Y, X | Y, Z | X, Z |
| Квадрат коэффициента регрессии  (b2) | 0,03 | 69059,13 | 2091813,91 |
| Квадрат парного коэффициента корреляции  (rx,y2) | 0,7633 | 0,9788 | 0,7901 |

Рассчитаем случайные ошибки коэффициента корреляции по формуле 3.13



, , (3.13)

где r - парный коэффициент корреляции;

n – количество предприятий в выборке.

Расчеты приведём в таблице 3.7.

Таблица 3.7

Случайные ошибки коэффициента корреляции

|  |  |  |  |
| --- | --- | --- | --- |
| Пары показателей | Mr (y, x) | Mr (y, z) | Mr(x, z) |
| Случайные ошибки коэффициента корреляции | 0,1538 | 0,0460 | 0,1449 |

Все необходимые расчёты были сделаны, можно найти коэффициент Стьюдента.

Таблица 3.8

Коэффициент Стьюдента парных коэффициентов корреляции

|  |  |  |  |
| --- | --- | --- | --- |
| Пары показателей | t (y, x) | t (y, z) | t (x, z) |
| Коэффициент Стьюдента | 4,9619 | 21,2655 | 5,4532 |

Сравним полученные результаты с табличным значением коэффициента Стьюдента. V= 12-2, вероятность a примем за 0,05. Табличное значение коэффициента равно 2.2281..

В паре показателей:

- зависимость объёмов добычи от стоимости основных фондов (X, Y) t(x,y)> tтаб, 4,9619>2,2281. Значит мы можем отклонить гипотезу о том, что r(x,y) в действительности равен 0, и лишь в силу случайных обстоятельств оказался равным проверяемой величине;

- зависимость объёмов добычи от производительности труды (Z, Y) t(z,y)> tтаб, 21,2655>2,2281. Значит мы можем отклонить гипотезу о том, что r(z,y) в действительности равен 0, и лишь в силу случайных обстоятельств оказался равным проверяемой величине;

- зависимость стоимости основных фондов от производительности (X, Z) t(x,z)> tтаб , 5,4532>2,2281. Значит мы не можем отклонить гипотезу о том, что r(x,c) в действительности равен 0, и лишь в силу случайных обстоятельств оказался равным проверяемой величине.

Коэффициенты корреляции во всех значимые. А корреляционно-регрессионная модель адекватная.

ЗАКЛЮЧЕНИЕ

Наиболее сложным этапом, завершающим регрессионный анализ, является интерпретация полученных результатов, т.е. перевод их с языка статистики и математики на язык экономики.

Интерпретация моделей регрессии осуществляется методами той отрасли знаний, к которой относятся исследуемые явления. Всякая интерпретация начинается со статистической оценки уравнения регрессии в целом и оценки значимости входящих в модель факторных признаков, т.е. с изучения, как они влияют на величину результативного признака. Чем больше величина коэффициента регрессии, тем значительнее влияние данного признака на моделируемую обработку биржевых ставок. Особое значение при этом имеет знак перед коэффициентом регрессии. Знаки коэффициентов регрессии говорят о характере влияния на результативный признак статистической обработки биржевых ставок. Если факторный признак имеет плюс, то с увеличением данного фактора результативный признак возрастает; если факторный признак со знаком минус, то с его увеличением результативный признак уменьшается. Интерпретация этих знаков полностью определяется социально-экономическим содержанием моделируемого признака. Если его величина изменяется в сторону увеличения, то плюсовые знаки факторных признаков имеют положительное влияние. При изменении результативного признака в сторону снижения положительные значения имеют минусовые знаки факторных признаков. Если экономическая теория подсказывает, что факторный признак должен иметь положительное значение, а он со знаком минус, то необходимо проверить расчеты параметров уравнения регрессии.

Корреляционный и регрессионный анализ позволяет определить зависимость между факторами, а так же проследить влияние задействованных факторов. Эти показатели имеют широкое применение в обработке статистических данных для достижения наилучших показателей биржевых ставок.[2]

СПИСОК ЛИТЕРАТУРЫ

1. Экономика: Учебник 3-е издание Под ред. А. С. Булатова. – М.: Юристъ, 2002. – 862с.

2. Экономика предприятия: Учебник/Под ред. проф. О.И.Волкова.- 2-е изд., М.: ИНФРА – М, 2001.- 520 с.–(Серия «Высшее образование»).

3. Статистика / А.И. Гинзбург. – СПб: Питер, 2002. – 128 с.

4. «Теория статистики» В.М. Гусаров, М.: ЮНИТИ, 2001. – 247 с.

5. Статистика: Учебник / И.И. Елисеева, Е.Б. Капралова; под ред. И.И. Елисеевой. – М.: КНОРУС, 2006. – 552 с. Финансы предприятий:

6. Статистика: Учебник / Под ред. И.И. Елисеевой. – М.: Высшее образование, 2006 – 565 с

7. Социально-экономическая статистика: Учеб. пособие / И.И. Колесникова. – Мн.: Новое издание, 2002. – 250 с.

8. Учебник / Колчина Н.В., Поляк Г.Б., Павлова Л.П. и др., под ред. Колчиной Л.В. – 2-е изд., перераб. и доп. – М.: ЮНИТИ-ДАНА, 2002. – с. 413

9. Николаева С.А. и др. Бухгалтерский учет основных средств: Справочное и учебно-методическое пособие. - Москва: Аналитика-Пресс, 2001.

10. Савицкая Г.В. Анализ хозяйственной деятельности: 5-е издание, переработанное и дополненное. Мн.: ИП «Экоперспектива», 2003. – с. 498..

11. Статистика: Учебное пособие / Харченко Л.П., Ионин В.Г. и др.; Под ред. В.Г. Ионина. – Изд. 2-е, М.: ИНФРА – М., 2006. – 384 с.

12. «Теория статистики», учебник под ред. Р.А. Шмойловой М.: Финансы и статистика, 2000. - 510 с.

13. «Практикум по теории статистики»., под ред. Р.А. Шмойловой, М.: Финансы и статистика, 2001. - 456 с.

14. Статистический словарь / Госкомстат РФ. – М.: Финансы и статистика, 1989. - 621 с.

15. Российский статистический ежегодник. 2004: Стат. сб. / Росстат. – М., 2004. – 725 с.