# **Содержание:**

Содержание: 1

Вступление 2

Теоретическая часть. 3

Многомерный корреляционный анализ 3

Многошаговый регрессионный анализ. 4

Многомерный регрессионный анализ 5

Метод отсева факторов по t-критерию 9

Практическая часть. 10

Вариационные характеристики. 10

Корреляционный анализ. 14

Многомерный регрессионный анализ. 15

Многошаговый регрессионный анализ. 16

**Начальный корреляционный анализ.** 17

Приложение: Олимп курсовая итог. 21

Использованная литература: 30

## Вступление

Для достоверного отображения объективно существующих в экономике процессов необходимо выявить существенные взаимосвязи между ними. В естественных науках часто речь идет о функциональной связи, когда каждому значению одной переменной соответствует вполне определенной значение другой. В экономике в большинстве случаев между переменными величинами существуют зависимости, когда каждому значению одной переменной соответствует не какое-то определенное, а множество возможных значений другой переменной. Такая зависимость получила название стохастической.

Частными случаями стохастической связи являются корреляционная и регрессионная связи.

Две случайные величины имеют корреляционную связь, если математическое ожидание одной из них изменяется в зависимости от изменения другой. Метод математической статистики, изучающий корреляционные связи между явлениями, называется корреляционным анализом. Основной его задачей является выявление связи между случайными переменными и оценка ее тесноты.

Но не все факторы, влияющие на экономические процессы, являются случайными величинами. Поэтому при анализе экономических явлений обычно рассматриваются связи между случайными и неслучайными величинами. Такие связи называются регрессионными, а метод математической статистики, их изучающий, называется регрессионным анализом. Кроме того, при изучении экономических процессов необходимо не только выявить связь между переменными, но и изучить и установить ее форму, что и является основной задачей регрессионного анализа.

Поэтому, как видно из написанного выше, многомерный регрессионный анализ, изучению экономических процессов с помощью которого и посвящена настоящая работа, будет гораздо подробнее и точнее при включении в него необходимых элементов корреляционного анализа.

# ***Теоретическая часть.***

## Многомерный корреляционный анализ

В многомерной модели корреляционного анализа (с четырьмя и более переменными) вычисление частных и множественных коэффициентов корреляции основывается на использовании матрицы коэффициентов парной корреляции.

Порядок частного коэффициента корреляции определяется количеством фиксируемых переменных. Выборочный частный коэффициент корреляции любого порядка можно определить по формуле



Это выражение предполагает вычисление большого числа выборочных частных коэффициентов корреляции от нулевого до (*к*-3)-го порядка, что является достаточно трудоемкой операцией.

Более удобным является вычисление частных коэффициентов корреляции по следующей схеме.

На основе матрицы выборочных коэффициентов парной корреляции

 (1)

где Q – симметричная положительно определенная матрица, имеем

 (2)

 (3)

и так далее, где

Dij – определитель матрицы, образованной из матрицы (1) вычеркиванием i-ой строки и j-го столбца для каждого определителя соответственно.

Для проверки значимости частного коэффициента корреляции используется величина t, имеющая t-распределение Стьюдента с числом степеней свободы =n-l-2:

, (4)

где n – число наблюдений;

l – число фиксированных переменных;

rчаст – соответствующий выборочный частный коэффициент корреляции.

С помощью таблицы распределения Стьюдента по уровню значимости α и =n-l-2 находится tкр. При tн >tкр гипотеза Но:ρчаст = 0 отвергается.

Доверительный интервал для частных коэффициентов корреляции строится при помощи z-преобразования Фишера

, аналогично рассмотренным ранее случаям.

Для определения тесноты связи между зависимой переменной и совокупностью объясняющих переменных используется выборочный коэффициент множественной корреляции, определяемый по формуле

, (5)

где D – определитель матрицы выборочных коэффициентов корреляции;

Dii – алгеброическое дополнение к элементу rii.

Для проверки значимости коэффициента множественной корреляции используется величина

, (6)

имеющая F-распределение с 1=l и =n-l-2 степенями свободы.

## Многошаговый регрессионный анализ.

Очевидно, что простое поверхностное изучение данных не позволяет обнаружить, какие факторы, рассмотренные на стадии статистического анализа исходной информации, являются существенными, а какие – нет. Может случиться, что якобы отсутствующая корреляция с данным фактором обнаруживается после того, как связь с другим фактором уже исключена.

Необходимо найти оптимальный вариант модели, отражающий основные закономерности исследуемого явления с достаточной степенью статистической надежности.

В модель должны быть включены все факторы, которые с экономической точки зрения оказывают влияние на зависимую переменную (в нашем случае – средняя продолжительность жизни). При невыполнении этого требования модель может оказаться неадекватной вследствие недоучета существенных факторов.

С другой стороны, количество факторов, включаемых в модель, не должно быть слишком большим. Невыполнение этого требования приводит к необходимости увеличения числа наблюдений, к невозможности использования достаточно сложных зависимостей, к снижению точности оценок, к сложности интерпретации модели и к трудности ее практического использования.

Таким образом, возникает задача уменьшения числа переменных, включаемых в модель, без нарушения исходных предпосылок, т.е. задача понижения размерности модели.

Выделяют два существенных подхода к решению проблемы сокращения количества исходных переменных:

1. отсеивание менее существенных факторов в процессе построения регрессионной модели;
2. замена исходного набора переменных меньшим числом эквивалентных факторов, полученных в результате преобразований исходного набора.

Процедура отсева несущественных факторов в процессе построения регрессионной модели и получила название многошагового регрессионного анализа.

Этот метод основан на вычислении нескольких промежуточных уравнений регрессии, в результате анализа которых получают конечную модель, включающую только факторы, оказывающие статистически существенное влияние на исследуемую зависимую переменную. Различные сочетания одних и тех же факторов оказывают разное влияние на зависимую переменную. Вследствие этого появляется необходимость выбора наилучшей модели, т.к. перебирать все возможные варианты сочетания факторов и строить множество уравнений регрессии (количество которых может быть очень велико) просто не имеет смысла.

Таким образом методы пошагового регрессионного анализа позволяют избежать столь громоздких расчетов и получить достаточно надежную и полную модель зависимости исследуемого признака от ряда объясняющих переменных.

Как было сказано выше, основой многошагового регрессионного анализа является построение уравнения регрессии. Рассмотрим более подробно его систему и основные понятия.

## Многомерный регрессионный анализ

В общем виде многомерная линейная регрессионная модель зависимости y от объясняющих переменных , ,…, имеет вид:

.

Для оценки неизвестных параметров  взята случайная выборка объема n из (k+1)–мерной случайной величины (y, ,,…,).

В матричной форме модель имеет вид:

,

где  , , , ε=

- вектор-столбец фактических значений зависимой переменной размерности n;

- матрица значений объясняющих переменных размерности n\*(k+1);

- вектор-столбец неизвестных параметров, подлежащих оценке, размерности (k+1);

- вектор-столбец случайных ошибок размерности n с математическим ожиданием ME=0 и ковариационной матрицей  соответственно, при этом

 -единичная матрица размерности (nxn).

Оценки неизвестных параметров  находятся методом наименьших квадратов, минимизируя скалярную сумму квадратов  по компонентам вектора β.

Далее подставив выражение

 в ,

получаем скалярную сумму квадратов 

Условием обращения полученной суммы в минимум является система нормальных уравнений:

, (j=0,1,2,…,k) .

В результате дифференцирования получается:

.

При замене вектора неизвестных параметров β на оценки, полученные методом наименьших квадратов, получаем следующее выражение:

.

Далее умножив обе части уравнения слева на матрицу , получим



Так как  , тогда .

Полученные оценки вектора b являются не смещенными и эффективными.

Ковариационная матрица вектора b имеет вид:

, где  - остаточная дисперсия.

Элементы главной диагонали этой матрицы представляют собой дисперсии вектора оценок b. Остальные элементы являются значениями коэффициентов ковариации:

, где  , .

Таким образом, оценка  - это линейная функция от зависимой переменной. Она имеет нормальное распределение с математическим ожиданием  и дисперсией .

Несмещенная оценка остаточной дисперсии определяется по формуле:

, где n – объем выборочной совокупности;

k – число объясняющих переменных.

Для проверки значимости уравнения регрессии используют F-критерий дисперсионного анализа, основанного на разложении общей суммы квадратов отклонений на составляющие части:

 , где  - сумма квадратов отклонений (от нуля), обусловленная регрессией;

 - сумма квадратов отклонений фактических значений зависимой переменной от расчетных , т.е. сумма квадратов отклонений относительно плоскости регрессии, обусловленное воздействием случайных и неучтенных в модели факторов.

Для проверки гипотезы  используется величина , которая имеет F-распределение Фишера-Снедекора с числом степеней свободы  и . Если  , то уравнение регрессии значимо, т.е. в уравнении есть хотя бы один коэффициент регрессии, отличный от нуля.

В случае значимости уравнения регрессии проверяется значимость отдельных коэффициентов регрессии. Для проверки нулевой гипотезы  используется величина

 , которая имеет F-распределение Фишера-Снедекора с числом степеней свободы  и ;  - соответствующий элемент главной диагонали ковариационной матрицы.

Коэффициент регрессии  считается значимым, если . Для значимых коэффициентов регрессии можно построить доверительные интервалы, используя формулу

 , где  находится по таблице распределения Стьюдента для уровня значимости  и числа степеней свободы .

В многошаговом регрессионном анализе наиболее известны три подхода:

1. Метод случайного поиска с адаптацией. Осуществляется путем построения нескольких уравнений регрессии на основе формально разработанного принципа включения факторов и последующего выбора лучшего уравнения с точки зрения определенного критерия.
2. Метод включения переменных, основанный на построении уравнения регрессии по одному значимому фактору и последовательном добавлении всех остальных статистически значимых переменных путем расчета частных коэффициентов корреляции и F-критерия при проверке значимости вводимого в модель фактора.
3. Метод отсева факторов по t-критерию. Данный метод заключается в построении уравнений регрессии по максимально возможному количеству объясняющих переменных и последующем исключении статистически не существенных факторов.

## Метод отсева факторов по t-критерию

Наиболее оправданным является использование многошагового регрессионного анализа, основанного на оценке значимости коэффициентов регрессии с помощью t-критерия Стьюдента. Данный метод и был использован при анализе продолжительности жизни населения стран Африки в данной курсовой работе, потому что его применение четко формализовано, и в то же время на различных стадиях построения модели можно производить качественный экономический анализ. Рассмотрим его более подробно.

Итак, на первом этапе строится уравнение регрессии по переменным, предположительно влияющим на исследуемую зависимую переменную. Затем с помощью определенных критериев исключаются те переменные, которые оказывают статистически несущественное влияние. На этом подходе основан метод отсева факторов по t-критерию в многошаговом регрессионном анализе.

Применение t-критерия при отборе существенных факторов основано на следующей предпосылке регрессионного анализа: если выполняется условие, что *Ei* распределены нормально, то величина  распределена по закону Стьюдента с *ν = n-k-1* степенями свободы. По этому критерию можно проверить гипотезу о существенном отличии от нуля коэффициента регрессии *βj*  при некотором заданном уровне значимости и *n-k-1* степенях, то коэффициент регрессии *βj*  признается значимым.

Простейшая схема проверки сводится к построению доверительного интервала для каждого коэффициента регрессии и проверке гипотезы о том, находится ли нуль внутри построенного интервала. Если это так, то данный коэффициент регрессии признается незначимым или же его значимость подвергается сомнению и выявляется на следующих этапах анализа.

Схема отбора значимых факторов в уравнение регрессии с помощью t-критерия выглядит следующим образом. Если все коэффициенты регрессии значимы, то уравнение регрессии признается окончательным и принимается в качестве модели исследуемого признака для последующего анализа. Если же среди коэффициентов регрессии имеются незначимые, то соответствующие объясняющие переменные следует исключить из уравнения.

Однако предварительно следует проранжировать коэффициенты регрессии по величине tH и в первую очередь отсеять тот фактор, для которого коэффициент регрессии незначим и tH  имеет наименьшее значение. Затем уравнение регрессии пересчитывается снова (уже без исключенного фактора), и производится оценка коэффициентов регрессии по t-критерию. Такую процедуру повторяют до тех пор, пока все коэффициенты регрессии в уравнении не окажутся значимыми.

При этом на каждом шаге, кроме формальной статистической проверки значимости коэффициентов регрессии, проводится экономический анализ несущественных факторов и устанавливается порядок их исключения. В некоторых случаях значение tH находится вблизи tкр, и, с точки зрения содержательности модели, этот фактор можно оставить для последующей проверки его значимости в сочетании с другим набором факторов. Возможность такого экономического анализа при формальной статистической процедуре отсеивания незначимых факторов по t-критерию является большим преимуществом этого метода многошагового регрессионного анализа.

Вместе с тем следует отметить, что несущественность коэффициента регрессии по t-критерию не всегда является надежным основанием для исключения переменной из дальнейшего анализа. Поэтому в ряде случаев для проведения многошагового регрессионного анализа с помощью t-критерия предполагается использовать некоторые дополнительные эмпирические процедуры. Например, исключать переменную из уравнения регрессии лишь в том случае, когда средняя квадратическая ошибка коэффициента регрессии превышает абсолютный размер вычисленного коэффициента, то есть когда tH по абсолютной величине меньше единицы. При этом предполагается, что нет достаточных логических оснований для того, чтобы оставлять такую переменную в модели.

# **Практическая часть.**

## Вариационные характеристики.

Для изучения корреляционного и регрессионного анализа в более подробном разрезе была взята совокупность стран Африки.

Задачей проводимого исследования является выявление и изучение зависимости данных экономических явлений.

При проведении данного исследования была взята совокупность, состоящая из 25 стран Африки:

Алжир, Ангола, Генин, Ботсвана, Бурунди, Буркина Фасо, Габон, Гамбия, Гана, Гвинея, Гвинея-Бисау, Джибути, Египет, Заир, Замбия, Зимбабве, Кабо-Верде, Кения, Коморские острова, Конго, Кот-д’Ивуар, Лесото, Либерия, Ливия.

Характеризующими являются следующие признаки: средняя продолжительность жизни (лет), численность населения (тыс. человек), доля городского населения (%), число медицинских работников на 10 тысяч населения (чел.), доля неграмотных (%), среднегодовой индекс роста производства продовольствия (%).

Но для упрощения проведения расчетов и всего исследования, а также выявления связи стоит разделить вышеописанные признаки на факторный и результативные и заменить их условными переменными (у, х1, х2, х3, х4, х5):

результативный признак (у) представляет собой среднюю продолжительность жизни (лет);

факторные признаки (х):

х1: численность населения (тыс. человек);

х2: доля городского населения (%);

х3: число медицинских работников на 10 тысяч населения (чел.);

х4: доля неграмотных (%);

х5: среднегодовой индекс роста производства продовольствия (%).

Начальные данные представлены в таблице:

┌────┬────────┬───────────┬────────┬────────┬────────┬─────────┐

│ N │ y │ x1 │ x2 │ x3 │ x4 │ x5 │

├────┼────────┼───────────┼────────┼────────┼────────┼─────────┤

│ 1 │ 63.00 │ 23102.00 │ 60.85 │ 32.70 │ 55.30 │ 87.00 │

│ 2 │ 44.50 │ 9226.00 │ 21.00 │ 12.70 │ 97.00 │ 58.00 │

│ 3 │ 46.00 │ 4304.00 │ 30.80 │ 7.50 │ 75.20 │ 108.00 │

│ 4 │ 56.50 │ 1169.00 │ 29.50 │ 35.80 │ 59.30 │ 71.00 │

│ 5 │ 48.50 │ 5001.00 │ 2.29 │ 3.80 │ 77.40 │ 101.00 │

│ 6 │ 47.20 │ 8305.00 │ 8.48 │ 8.10 │ 91.20 │ 92.00 │

│ 7 │ 51.00 │ 1058.00 │ 35.80 │ 22.30 │ 87.60 │ 98.00 │

│ 8 │ 37.00 │ 670.00 │ 18.50 │ 15.10 │ 85.20 │ 62.00 │

│ 9 │ 54.00 │ 13704.00 │ 35.86 │ 37.60 │ 69.80 │ 73.00 │

│ 10 │ 42.20 │ 6380.00 │ 19.07 │ 4.20 │ 80.00 │ 91.00 │

│ 11 │ 45.00 │ 925.00 │ 23.80 │ 38.60 │ 71.60 │ 83.00 │

│ 12 │ 64.50 │ 372.00 │ 73.95 │ 72.20 │ 80.00 │ 75.00 │

│ 13 │ 60.60 │ 50740.00 │ 45.37 │ 47.90 │ 56.50 │ 89.00 │

│ 14 │ 52.00 │ 32461.00 │ 39.50 │ 12.60 │ 42.10 │ 86.00 │

│ 15 │ 53.30 │ 7563.00 │ 40.40 │ 18.50 │ 56.00 │ 91.00 │

│ 16 │ 57.80 │ 8640.00 │ 19.60 │ 16.60 │ 29.20 │ 94.00 │

│ 17 │ 53.00 │ 10822.00 │ 34.60 │ 14.40 │ 59.50 │ 102.00 │

│ 18 │ 61.50 │ 348.00 │ 5.80 │ 18.80 │ 63.10 │ 83.00 │

│ 19 │ 53.30 │ 22936.00 │ 14.17 │ 11.20 │ 50.40 │ 93.00 │

│ 20 │ 52.00 │ 472.00 │ 11.53 │ 15.30 │ 41.60 │ 91.00 │

│ 21 │ 48.50 │ 1837.00 │ 37.27 │ 31.70 │ 84.40 │ 83.00 │

│ 22 │ 52.30 │ 11142.00 │ 37.62 │ 13.50 │ 58.80 │ 102.00 │

│ 23 │ 50.60 │ 1619.00 │ 4.52 │ 0.50 │ 48.00 │ 78.00 │

│ 24 │ 51.00 │ 2349.00 │ 32.94 │ 11.30 │ 74.60 │ 91.00 │

│ 25 │ 60.80 │ 4083.00 │ 52.40 │ 64.80 │ 49.90 │ 151.00 │

└────┴────────┴───────────┴────────┴────────┴────────┴─────────┘

Реализация алгоритма многомерного регрессионного анализа начинается с расчета важнейших статистических характеристик исходной информации и матрицы выборочных парных коэффициентов корреляции.

Рассмотрим более подробно вариационные характеристики переменной у:

. число наблюдений 25

. среднее значение 52.2440

. верхняя оценка среднего 54.5134

. нижняя оценка среднего 49.9746

. среднеквадратическое отклонение 6.6138

. дисперсия 43.7425

. дисперсия (несмещ. оценка) 45.5651

. среднекв. откл. (несмещ. оценка) 6.7502

. среднее линейное отклонение 5.0938

. моменты начальные

. 2-го поpядка 2773.1780

. 3-го поpядка 1.4943e+05

. 4-го поpядка 8.1668e+06

. моменты центpальные

. 3-го поpядка -2.1613e+01

. 4-го поpядка 5.1166e+03

. коэффициент асимметрии

. значение -0.0747

. несмещенная оценка -0.0796

. среднекв. отклонение 0.4637

. коэффициент эксцесса

. значение -0.0000

. несмещенная оценка 0.2846

. среднекв. отклонение 0.9017

. коэффициенты вариации

. по pазмаху 0.5264

. сpеднему линейному откл. 0.0975

. сpеднеквадp. откл. 0.1266

. медиана 52.0000

. мода 48.5000

. минимальное значение 37.0000

. максимальное значение 64.5000

. размах 27.5000

Проанализируем их.

Средняя продолжительность жизни в странах Африки – 52,244 года. Она вычисляется по формуле средней арифметической невзвешенной:

\_

у = Σуi/n

где n – объем исследуемой совокупности.

Дисперсия в нашем случае равна 43,7425. Она представляет собой средний квадрат отклонений индивидуальных значений признака от их средней величины и вычисляется по формуле:

\_

σ2 = Σ (у I – у )2 / n

Среднее квадратическое отклонение представляет собой корень второй степени из дисперсии, и в нашем случае σ = 6,6138, то есть значение продолжительности жизни в среднем отклоняется на 6,6138 лет.

А среднее линейное отклонение вычисляется по формуле:

\_ \_

d = Σ |уi -y| / n,

которое в нашем случае равно 5,0938 и представляет собой среднюю величину из отклонений вариантов признака от их средней.

Коэффициент вариации среднеквадратического отклонения в исследуемой нами совокупности равен Vσ = 0,1266 или 12,66%, который вычисляется по формуле:

\_

Vσ = σ / у \* 100%.

Коэффициент вариации характеризует не только сравнительную оценку вариации, но и дает характеристику однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33%, то есть наша совокупность является однородной.

Мода – значение признака, наиболее часто встречающегося в совокупности. Она рассчитывается по формуле:

Мо = уМо  + iМо \* (fМо – fМо-1)/(fМо – fМо-1)\*(fМо – fМо+1)

То есть по Африке наиболее часто встречающееся значение продолжительности жизни равно 48,5 лет.

Медиана – значение признака, приходящегося на середину ранжированной (упорядоченной) совокупности.

Ме = уМе  + iМе \* (0,5 Σf – SМе-1)/fМе.

Таким образом, в нашем случае в половине стран Африки население имеет среднюю продолжительность жизни менее 52 лет, а в другой половине – более 52 лет.

Начальным моментом порядка k случайной величины х называют математическое ожидание величины хк:

νк  = М (хк),

в частности ν1 = М (х), ν2  = М (х2).

В нашем случае

начальные моменты равны:

. 2-го поpядка 2773.1780

. 3-го поpядка 1.4943e+05

. 4-го поpядка 8.1668e+06

Центральным моментом порядка k случайной величины х называют математическое ожидание величины (х – (М (х))к, в частности

μ1 = М[х – М (х)] = 0; μ2 = М[ ( х – М (х))2] = D (х).

В нашем случае центральные моменты равны:

. 3-го поpядка -2.1613e+01

. 4-го поpядка 5.1166e+03

Теперь рассмотрим нашу совокупность на предмет симметрии.

Симметричным называется распределение, в котором частоты любых двух вариантов, равностоящих в обе стороны от центра распределения, равны между собой. В статистике для характеристики асимметрии используют показатели асимметрии и эксцесса.

Так как видно, что наша совокупность асимметричная, найдем степень асимметрии. Сперва используем коэффициент асимметрии:

\_

Аs = (у – Мо)/ σ = 0,4637,

что свидетельствует о наличии незначительной правосторонней асимметрии (Аs>0).

Теперь рассчитаем показатель эксцесса:

ЕК = μ4/ σ4 – 3, где μ4 – центральный момент четвертого порядка.

ЕК  = 0,9017, следовательно, распределение стран Африки по продолжительности жизни является островершинным (ЕК>0).

Кроме того, взглянув на нашу совокупность, можно увидеть, что максимальная продолжительность жизни жителей стран Африки равна уmax=64,5 лет, а минимальная у min=37 лет.

Размах данной совокупности равен уmax  - у min  = 27,5 лет.

## Многошаговый регрессионный анализ.

Построим корреляционную модель из исследуемых шести переменных:y,, ,,,.

Присвоим для облегчения обозначений всем переменным порядковые номера: у-1, х1-2, х2-3, x3-4,x4-5,x5-6.

Предварительно, с целью анализа взаимосвязи показателей построена таблица парных коэффициентов корреляции R.

┌─────┬───────┬───────┬───────┬───────┬───────┬───────┐

│ │ y │ x1 │ x2 │ x3 │ x4 │ x5 │

├─────┼───────┼───────┼───────┼───────┼───────┼───────┤

│ y │ 1.00 │ 0.30 │ 0.53 │ 0.60 │ -0.51 │ 0.26 │

│ x1 │ 0.30 │ 1.00 │ 0.27 │ 0.10 │ -0.33 │ 0.02 │

│ x2 │ 0.53 │ 0.27 │ 1.00 │ 0.74 │ -0.04 │ 0.17 │

│ x3 │ 0.60 │ 0.10 │ 0.74 │ 1.00 │ -0.03 │ 0.15 │

│ x4 │ -0.51 │ -0.33 │ -0.04 │ -0.03 │ 1.00 │ -0.31 │

│ x5 │ 0.26 │ 0.02 │ 0.17 │ 0.15 │ -0.31 │ 1.00 │

└─────┴───────┴───────┴───────┴───────┴───────┴───────┘

Анализ матрицы парных коэффициентов корреляции показывает, что результативный показатель наиболее тесно связан с показателем x3 – числом медицинских работников на 10 тысяч населения (ryx3=0.60).

Одним из основных препятствий эффективного применения регрессионного анализа, является мультиколлинеарность (наличие сильной корреляции между независимыми переменными, входящими в уравнение регрессии x1,x2,x3,x4,x5). Наиболее распространенный метод выявления коллинеарности основан на анализе парных коэффициентов корреляции. Он состоит в том, что две или несколько переменных признаются коллинеарными (мультиколлинеарными), если парные коэффициенты корреляции больше определенной величины. На практике наиболее часто считают, что два аргумента коллинеарны, если парный коэффициент корреляции между ними по абсолютной величине больше 0,8.

В данном примере ни один парный коэффициент корреляции не превышает величины 0,8, что говорит об отсутствии явления мультиколлинеарности.

Приступим непосредственно к регрессионному анализу.

Построим регрессионную модель по следующим факторам: х1, х2, х3, х4 и х5. Для расчета параметров уравнения регрессии используем стандартную программу многошагового регрессионного анализа с последовательным отсевом факторов.

На первом шаге построения модели в уравнение линейной регрессии вводятся все указанные выше переменные. В результате получена следующая модель:

*ŷ*= 57.700+0.000\*x1+0.056\*x2+0.173\*x3-0.182\*x4+0.007\*x5.

Прежде чем осуществлять проверку значимости уравнения регрессии и коэффициентов регрессии, следует убедиться, что выполняется необходимое для этого условие, а именно следует проверить, является ли распределение остатков (т.е. отклонений эмпирических значений зависимой переменной от расчетных) нормальным. Для проверки данного условия используем критерий согласия Пирсона , рассчитанные значения которого приведены ниже:

Проверка нормального закона распределения

критерий хи-квадpат

.число степеней свободы 3

.хи-квадpат pасчетное 1.571

веpоятн. хи-квадpат заключение

уpовень теоpетическое о гипотезе

0.900 6.226 не отвеpгается

0.950 7.795 не отвеpгается

0.990 11.387 не отвеpгается

Таким образом, можно сделать вывод, что гипотеза о нормальности распределения остатков не отвергается с доверительной вероятностью 0.95 (=7.795).

Проверка значимости уравнения регрессии показала, что оно значимо на уровне доверительной вероятности 0,95. (см. приложение 3.1)

Уровень множественного коэффициента детерминации (0,625) свидетельствует о том, что воздействием включенных в модель факторов обусловлено 62,5% вариации средней продолжительности жизни в странах Африки.

Далее осуществляется проверка значимости отдельных коэффициентов регрессии на основе t-критерия Стьюдента. Для определения , используем таблицу распределения Стьюдента: =2,093 (α=0,05 и ν=n-k-1=25-5-1=19).

По нижеприведенной таблице (гр.5 t-значения) статистически существенными оказались только два коэффициента регрессии при переменных  и  (|t|>).

Оценки коэффициентов линейной регрессии

┌───┬──────────┬───────────┬───────────────┬───────────┬────────┬─────────┐

│ N │ Значение │ Дисперсия │ Средне- │ t - │ Нижняя │ Верхняя │

│ │ │ │ квадатическое │ значение │ оценка │ оценка │

│ │ │ │ отклонение │ │ │ │

├───┼──────────┼───────────┼───────────────┼───────────┼────────┼─────────┤

│ │ 57.70 │ 59.12 │ 7.69 │ 7.50 │ 44.37 │ 71.03 │

│ │ 0.00 │ 0.00 │ 0.00 │ 0.36 │ -0.00 │ 0.00 │

│ │ 0.06 │ 0.01 │ 0.08 │ 0.66 │ -0.09 │ 0.20 │

│ │ 0.17 │ 0.01 │ 0.08 │ 2.21 │ 0.04 │ 0.31 │

│ │ -0.18 │ 0.00 │ 0.06 │ -2.96 │ -0.29 │ -0.08 │

│ │ 0.01 │ 0.00 │ 0.06 │ 0.12 │ -0.09 │ 0.11 │

└───┴──────────┴───────────┴───────────────┴───────────┴────────┴─────────┘

Среди незначимых коэффициентов регрессии наименее существенно по значению t-критерия является коэффициент регрессии при переменной  (среднегодовой индекс роста производства продовольствия), t=0.12. Этот фактор и подлежит исключению из модели в первую очередь. 

Исключив указанный фактор, на втором шаге получаем уравнение регрессии следующего вида:

*ŷ*= 58.478+0.000\*x1+0.057\*x2+0.173\*x3-0.184\*x4 .

Величина коэффициента детерминации на этом шаге не изменилась и составляет 0,625, гипотеза о значимости уравнения также не отвергается с вероятностью 0,95 (см. приложение 3.2).

Т.к. значение степеней свободы на каждом этапе построения модели изменяется (в связи с уменьшением числа объясняющих переменных), то  также меняется. Тогда при α=0,05 и

ν=n-k-1=25-4-1=20, =2,086. Таким образом, значимыми являются коэффициенты регрессии при факторах  и , а среди оставшихся незначимых наименьшее значение t-критерия, которое равно 0,35, принадлежит коэффициенту регрессии при переменной . Поэтому фактор  (численность населения) из дальнейшего процесса исключается.

На третьем шаге уравнение регрессии имеет следующий вид:

*ŷ*= 59.036+0.066\*x2+0.168\*x3-0.191\*x4 .

Воздействием включенных в модель переменных объясняется 62,2% вариации средней продолжительности жизни. Проверка на значимость уравнения регрессии показала, что оно значимо (на уровне значимости α=0,05). На этом шаге =2,080 (α=0,05 и ν=n-k-1=25-3-1=21), таким образом, статистически существенными оказались все коэффициенты регрессии, кроме коэффициента при объясняющей переменной , который и подлежит исключению по t-критерию из уравнения регрессии (t=0,87).

На последнем шаге регрессионного анализа получено значимое уравнение следующего вида:

Y=59.951+0.215x3-0.192x4.

Все коэффициенты регрессии значимы (см. приложение).

В результате моделирования зависимости средней продолжительности жизни в странах Африки можно сделать следующие выводы.

Уровень множественного коэффициента детерминации 0,609 свидетельствует о том, что 60,9% вариации зависимой переменной объясняется вариацией двух факторов:

x3 - число медицинских работников на 10 тыс. населения,

x4 - доля неграмотных.

Указанный уровень влияния достаточно высок, поэтому можно сделать вывод, что все факторы, оказывающие существенной влияние на среднюю продолжительность жизни, включены в модель, поскольку уровень остаточной вариации составляет 39.1%, объясняется воздействием случайных и неучтенных в модели факторов.

В рассматриваемом уравнении регрессии с изменением каждого фактора на одну единицу собственного измерения (при постоянном значении остальных факторов, вошедших в модель) зависимая переменная изменяется на соответствующий коэффициент регрессии βj отражает среднее приращение функции за счет единичного приращения j-го аргумента, независимое от изменения остальных учтенных в модели аргументов. Интерпретируемый таким образом коэффициент регрессии используется в экономико-статистическом анализе как средняя оценка эффективности влияния j-го аргумента на функцию.

Значение коэффициента регрессии βj зависит от принятых единиц измерения величин у и хj. Если единица измерения хj велика, то увеличение хj на единицу соответствует меньшее изменение среднего значения у, то есть βj мало. Если единица измерения у велика, то соответствующее изменение у выражается большим количеством единиц хj, следовательно, βj велико.

Анализируя полученную модель, можно сказать, что при увеличении числа медицинских работников на 1 человека средняя продолжительность жизни жителей стран Африки повышается в среднем на 0.215 лет; при увеличении доли неграмотных на 1% средняя продолжительность жизни уменьшится на 0.192 лет (обратная зависимость).

Однако с помощью коэффициентов регрессии нельзя сопоставить факторы по степени их влияния на зависимую переменную из-за различия единиц измерения и разной степени колеблемости. Поэтому для устранения таких различий при интерпретации применяется целая система показателей: средние частные коэффициенты эластичности, бета-коэффициенты или коэффициенты регрессии в стандартизированном масштабе и дельта-коэффициенты.

Средний частный коэффициенты эластичности рассчитывается по формуле:

\_ \_

Эj = bj\*xj / y.

\_

В рассматриваемой модели при изменении на 1% числа медицинских работников на 10 тысяч населения и доли неграмотных среди жителей исследуемых стран Африки средняя продолжительность жизни изменяется следующим образом: увеличивается на 0.094% и уменьшается на 0.241% соответственно (частные коэффициенты эластичности). - см. приложение.

Однако средний частный коэффициент эластичности не учитывает степени колеблемости факторов, которая может значительно различаться у отдельных факторов. Поэтому для устранения различий в измерении и степени колеблемости факторов используется другой показатель - коэффициент регрессии в стандартизированном масштабе (бета-коэффициент). Он показывает, на какую часть величины среднего квадратического отклонения изменяется среднее значение зависимой переменной с изменением соответствующей независимой переменной на одно среднее квадратическое отклонение при фиксированном на постоянном уровне значении остальных независимых переменных.

Бета-коэффициенты, рассчитанные для нашей модели, показывают, что при увеличении на одно среднее квадратическое отклонение числа медработников на 10 тысяч населения и доли неграмотных, средняя продолжительность жизни в среднем увеличивается на 0.587 и уменьшается на 0.495 средних квадратических отклонений соответственно. - см. приложение.

С помощью частных коэффициентов эластичности и с помощью бета-коэффициентов можно проранжировать факторы по степени их влияния на зависимую переменную, то есть сопоставить их между собой по величине этого влияния. Но с помощью бета-коэффициентов нельзя непосредственно оценить долю влияния каждого фактора в суммарном влиянии всех факторов. Для этой цели используются дельта-кэффициенты.

В практических задачах при корректно проведенном анализе величины дельта-коэффициентов положительны, то есть все коэффициенты регрессии имеют тот же знак, что и соответствующие парные коэффициенты корреляции. В этих случаях сумма величин вкладов независимых переменных равна коэффициенту множественной детерминации. Вместе с тем, в некоторых исследованиях отдельные коэффициенты регрессии имеют знак, противоположный знаку соответствующего коэффициента парной корреляции, вследствие чего величина дельта-коэффициента будет отрицательной. Не менее важно, что случаи с отрицательными вкладами могут иметь место только при значительной коррелированности объясняющих переменных.

В нашей модели наибольшее влияние на среднюю продолжительность жизни оказывает число медработников на 10 тысяч населения - 58.2%, а доля неграмотных оказывает влияние в размере 41.8%.

В настоящей курсовой работе был рассмотрен один из наиболее популярных в настоящее время методов математико-статистического моделирования экономических процессов, который позволяет строить достаточно адекватные и легко экономически интерпретируемые модели. Но легко заметить, что все вышеприведенные вычисление очень трудоемки и занимают немало времени. Поэтому, кроме вычислений вручную, а также для упрощения исследования, была проведена работа в пакете прикладных программ «ОЛИМП» - совокупность программных средств, ориентированных на решение задач экономического анализа и прогнозирования с помощью различных методов математической статистики. Полученные результаты приведены в Приложении.

# ***Приложение.***

Просмотр начальных данных

┌────┬────────┬───────────┬────────┬────────┬────────┬─────────┐

│ N │ y │ x1 │ x2 │ x3 │ x4 │ x5 │

├────┼────────┼───────────┼────────┼────────┼────────┼─────────┤

│ 1 │ 63.00 │ 23102.00 │ 60.85 │ 32.70 │ 55.30 │ 87.00 │

│ 2 │ 44.50 │ 9226.00 │ 21.00 │ 12.70 │ 97.00 │ 58.00 │

│ 3 │ 46.00 │ 4304.00 │ 30.80 │ 7.50 │ 75.20 │ 108.00 │

│ 4 │ 56.50 │ 1169.00 │ 29.50 │ 35.80 │ 59.30 │ 71.00 │

│ 5 │ 48.50 │ 5001.00 │ 2.29 │ 3.80 │ 77.40 │ 101.00 │

│ 6 │ 47.20 │ 8305.00 │ 8.48 │ 8.10 │ 91.20 │ 92.00 │

│ 7 │ 51.00 │ 1058.00 │ 35.80 │ 22.30 │ 87.60 │ 98.00 │

│ 8 │ 37.00 │ 670.00 │ 18.50 │ 15.10 │ 85.20 │ 62.00 │

│ 9 │ 54.00 │ 13704.00 │ 35.86 │ 37.60 │ 69.80 │ 73.00 │

│ 10 │ 42.20 │ 6380.00 │ 19.07 │ 4.20 │ 80.00 │ 91.00 │

│ 11 │ 45.00 │ 925.00 │ 23.80 │ 38.60 │ 71.60 │ 83.00 │

│ 12 │ 64.50 │ 372.00 │ 73.95 │ 72.20 │ 80.00 │ 75.00 │

│ 13 │ 60.60 │ 50740.00 │ 45.37 │ 47.90 │ 56.50 │ 89.00 │

│ 14 │ 52.00 │ 32461.00 │ 39.50 │ 12.60 │ 42.10 │ 86.00 │

│ 15 │ 53.30 │ 7563.00 │ 40.40 │ 18.50 │ 56.00 │ 91.00 │

│ 16 │ 57.80 │ 8640.00 │ 19.60 │ 16.60 │ 29.20 │ 94.00 │

│ 17 │ 53.00 │ 10822.00 │ 34.60 │ 14.40 │ 59.50 │ 102.00 │

│ 18 │ 61.50 │ 348.00 │ 5.80 │ 18.80 │ 63.10 │ 83.00 │

│ 19 │ 53.30 │ 22936.00 │ 14.17 │ 11.20 │ 50.40 │ 93.00 │

│ 20 │ 52.00 │ 472.00 │ 11.53 │ 15.30 │ 41.60 │ 91.00 │

│ 21 │ 48.50 │ 1837.00 │ 37.27 │ 31.70 │ 84.40 │ 83.00 │

│ 22 │ 52.30 │ 11142.00 │ 37.62 │ 13.50 │ 58.80 │ 102.00 │

│ 23 │ 50.60 │ 1619.00 │ 4.52 │ 0.50 │ 48.00 │ 78.00 │

│ 24 │ 51.00 │ 2349.00 │ 32.94 │ 11.30 │ 74.60 │ 91.00 │

│ 25 │ 60.80 │ 4083.00 │ 52.40 │ 64.80 │ 49.90 │ 151.00 │

└────┴────────┴───────────┴────────┴────────┴────────┴─────────┘

\*\*\* Вариационные характеристики переменной y \*\*\*

. число наблюдений 25

. среднее значение 52.2440

. верхняя оценка среднего 54.5134

. нижняя оценка среднего 49.9746

. среднеквадратическое отклонение 6.6138

. дисперсия 43.7425

. дисперсия (несмещ. оценка) 45.5651

. среднекв. откл. (несмещ. оценка) 6.7502

. среднее линейное отклонение 5.0938

. моменты начальные

. 2-го поpядка 2773.1780

. 3-го поpядка 1.4943e+05

. 4-го поpядка 8.1668e+06

. моменты центpальные

. 3-го поpядка -2.1613e+01

. 4-го поpядка 5.1166e+03

. коэффициент асимметрии

. значение -0.0747

. несмещенная оценка -0.0796

. среднекв. отклонение 0.4637

. коэффициент эксцесса

. значение -0.0000

. несмещенная оценка 0.2846

. среднекв. отклонение 0.9017

. коэффициенты вариации

. по pазмаху 0.5264

. сpеднему линейному откл. 0.0975

. сpеднеквадp. откл. 0.1266

. медиана 52.0000

. мода 48.5000

. минимальное значение 37.0000

. максимальное значение 64.5000

. размах 27.5000

\*\*\*\* Характеристики интеpвального pяда \*\*\*\*\*

. среднее значение 52.4000

. среднеквадратическое отклонение 6.5949

. дисперсия 43.4928

. коэффициент асимметpии -0.0815

. коэффициент эксцесса -0.2092

. медиана 51.5139

. мода 50.7500

N инт. Начало Сеpедина Конец Частота Частость

1 34.7083 37.0000 39.2917 1.0 0.0400

2 39.2917 41.5833 43.8750 1.0 0.0400

3 43.8750 46.1667 48.4583 4.0 0.1600

4 48.4583 50.7500 53.0417 9.0 0.3600

5 53.0417 55.3333 57.6250 4.0 0.1600

6 57.6250 59.9167 62.2083 4.0 0.1600

7 62.2083 64.5000 66.7917 2.0 0.0800

Пpовеpка ноpмального закона pаспpеделения

Кpитеpий хи-квадpат

.число степеней свободы 3

.хи-квадpат pасчетное 1.571

веpоятн. хи-квадpат заключение

уpовень теоpетическое о гипотезе

0.900 6.226 не отвеpгается

0.950 7.795 не отвеpгается

0.990 11.387 не отвеpгается

222222222222222 ОТЧЕТ 2222222222222222222222222222222222

0,990 11,387 не отвергается

или

не отвергается с вероятностью 0,950

32

Матpица

┌─────┬───────┬───────┬───────┬───────┬───────┬───────┐

│ N │ 1 │ 2 │ 3 │ 4 │ 5 │ 6 │

├─────┼───────┼───────┼───────┼───────┼───────┼───────┤

│ y │ 1.00 │ 0.30 │ 0.53 │ 0.60 │ -0.51 │ 0.26 │

│ x1 │ 0.30 │ 1.00 │ 0.27 │ 0.10 │ -0.33 │ 0.02 │

│ x2 │ 0.53 │ 0.27 │ 1.00 │ 0.74 │ -0.04 │ 0.17 │

│ x3 │ 0.60 │ 0.10 │ 0.74 │ 1.00 │ -0.03 │ 0.15 │

│ x4 │ -0.51 │ -0.33 │ -0.04 │ -0.03 │ 1.00 │ -0.31 │

│ x5 │ 0.26 │ 0.02 │ 0.17 │ 0.15 │ -0.31 │ 1.00 │

└─────┴───────┴───────┴───────┴───────┴───────┴───────┘

33333333333 ОТЧЕТ 33333333333333333333

НАЧАЛО

\*\*\* Протокол множественной линейной регрессии \*\*\*

Зависимая переменная Y - y

Функция Y = +57.700+0.000\*x1+0.056\*x2+0.173\*x3-0.182\*x4+0.007\*x5

Оценки коэффициентов линейной регрессии

┌───┬──────────┬───────────┬───────────────┬───────────┬────────┬─────────┐

│ N │ Значение │ Дисперсия │ Средне- │ t - │ Нижняя │ Верхняя │

│ │ │ │ квадатическое │ значение │ оценка │ оценка │

│ │ │ │ отклонение │ │ │ │

├───┼──────────┼───────────┼───────────────┼───────────┼────────┼─────────┤

│ 1 │ 57.70 │ 59.12 │ 7.69 │ 7.50 │ 44.37 │ 71.03 │

│ 2 │ 0.00 │ 0.00 │ 0.00 │ 0.36 │ -0.00 │ 0.00 │

│ 3 │ 0.06 │ 0.01 │ 0.08 │ 0.66 │ -0.09 │ 0.20 │

│ 4 │ 0.17 │ 0.01 │ 0.08 │ 2.21 │ 0.04 │ 0.31 │

│ 5 │ -0.18 │ 0.00 │ 0.06 │ -2.96 │ -0.29 │ -0.08 │

│ 6 │ 0.01 │ 0.00 │ 0.06 │ 0.12 │ -0.09 │ 0.11 │

└───┴──────────┴───────────┴───────────────┴───────────┴────────┴─────────┘

Кpитические значения t-pаспpеделения

пpи 19 степенях свободы

веpоятность t-значение

0.900 1.331

0.950 1.734

* 1. 2.542

Оценки коэффициентов интерпретации линейной регрессии

╔════╤════════╤═════════╤═════════╗

║ N │Коэффиц.│Вета- │Дельта- ║

║ │эластичн│коэффиц. │коэффиц. ║

╠════╪════════╪═════════╪═════════╣

║1 │ +0.006│ +0.056│ +0.027║

║2 │ +0.031│ +0.147│ +0.125║

║3 │ +0.075│ +0.471│ +0.455║

║4 │ -0.229│ -0.469│ +0.386║

║5 │ +0.012│ +0.019│ +0.008║

╚════╧════════╧═════════╧═════════╝

Таблица остатков

┌────┬──────────────┬───────────┬────────────┬───────────────┐

│ N │ Эмпирическое │ Расчетное │ Ошибка │ Ошибка │

│ │ значение │ значение │ абсолютная │ относительная │

├────┼──────────────┼───────────┼────────────┼───────────────┤

│ 1 │ 63.00 │ 58.05 │ 4.95 │ 0.08 │

│ 2 │ 44.50 │ 44.14 │ 0.36 │ 0.01 │

│ 3 │ 46.00 │ 47.95 │ -1.95 │ -0.04 │

│ 4 │ 56.50 │ 55.30 │ 1.20 │ 0.02 │

│ 5 │ 48.50 │ 45.30 │ 3.20 │ 0.07 │

│ 6 │ 47.20 │ 43.92 │ 3.28 │ 0.07 │

│ 7 │ 51.00 │ 48.36 │ 2.64 │ 0.05 │

│ 8 │ 37.00 │ 46.32 │ -9.32 │ -0.25 │

│ 9 │ 54.00 │ 54.47 │ -0.47 │ -0.01 │

│ 10 │ 42.20 │ 45.80 │ -3.60 │ -0.09 │

│ 11 │ 45.00 │ 53.30 │ -8.30 │ -0.18 │

│ 12 │ 64.50 │ 60.30 │ 4.20 │ 0.07 │

│ 13 │ 60.60 │ 60.49 │ 0.11 │ 0.00 │

│ 14 │ 52.00 │ 56.08 │ -4.08 │ -0.08 │

│ 15 │ 53.30 │ 53.86 │ -0.56 │ -0.01 │

│ 16 │ 57.80 │ 57.30 │ 0.50 │ 0.01 │

│ 17 │ 53.00 │ 52.38 │ 0.62 │ 0.01 │

│ 18 │ 61.50 │ 50.41 │ 11.09 │ 0.18 │

│ 19 │ 53.30 │ 52.66 │ 0.64 │ 0.01 │

│ 20 │ 52.00 │ 54.09 │ -2.09 │ -0.04 │

│ 21 │ 48.50 │ 50.57 │ -2.07 │ -0.04 │

│ 22 │ 52.30 │ 52.53 │ -0.23 │ -0.00 │

│ 23 │ 50.60 │ 49.92 │ 0.68 │ 0.01 │

│ 24 │ 51.00 │ 48.66 │ 2.34 │ 0.05 │

│ 25 │ 60.80 │ 63.95 │ -3.15 │ -0.05 │

└────┴──────────────┴───────────┴────────────┴───────────────┘

Характеристики остатков

Среднее значение..................... -0.000

Оценка дисперсии..................... 16.4

Оценка приведенной дисперсии........ 21.6

Средний модуль остатков.............. 2.866

Относительная ошибка аппроксимации... 0.057

Критерий Дарбина-Уотсона............. 1.857

Коэффициент детерминации............. 0.625

F - значение ( n1 = 6, n2 = 19)... 532

Гипотеза о значимости уравнения

не отвергается с вероятностью 0.950

\*\*\* Протокол множественной линейной регрессии \*\*\*

Зависимая переменная Y - y

Функция Y = +58.478+0.000\*x1+0.057\*x2+0.173\*x3-0.184\*x4

Оценки коэффициентов линейной регрессии

┌───┬──────────┬───────────┬───────────────┬───────────┬────────┬─────────┐

│ N │ Значение │ Дисперсия │ Средне- │ t - │ Нижняя │ Верхняя │

│ │ │ │ квадатическое │ значение │ оценка │ оценка │

│ │ │ │ отклонение │ │ │ │

├───┼──────────┼───────────┼───────────────┼───────────┼────────┼─────────┤

│ 1 │ 58.48 │ 18.27 │ 4.27 │ 13.68 │ 51.08 │ 65.87 │

│ 2 │ 0.00 │ 0.00 │ 0.00 │ 0.35 │ -0.00 │ 0.00 │

│ 3 │ 0.06 │ 0.01 │ 0.08 │ 0.70 │ -0.08 │ 0.20 │

│ 4 │ 0.17 │ 0.01 │ 0.08 │ 2.26 │ 0.04 │ 0.30 │

│ 5 │ -0.18 │ 0.00 │ 0.06 │ -3.27 │ -0.28 │ -0.09 │

└───┴──────────┴───────────┴───────────────┴───────────┴────────┴─────────┘

Кpитические значения t-pаспpеделения

пpи 20 степенях свободы

веpоятность t-значение

0.900 1.328

0.950 1.730

0.990 2.531

Оценки коэффициентов интерпретации линейной регрессии

╔════╤════════╤═════════╤═════════╗

║ N │Коэффиц.│Вета- │Дельта- ║

║ │эластичн│коэффиц. │коэффиц. ║

╠════╪════════╪═════════╪═════════╣

║1 │ +0.005│ +0.054│ +0.025║

║2 │ +0.032│ +0.150│ +0.128║

║3 │ +0.075│ +0.471│ +0.455║

║4 │ -0.232│ -0.476│ +0.392║

╚════╧════════╧═════════╧═════════╝

Таблица остатков

┌────┬──────────────┬───────────┬────────────┬───────────────┐

│ N │ Эмпирическое │ Расчетное │ Ошибка │ Ошибка │

│ │ значение │ значение │ абсолютная │ относительная │

├────┼──────────────┼───────────┼────────────┼───────────────┤

│ 1 │ 63.00 │ 58.12 │ 4.88 │ 0.08 │

│ 2 │ 44.50 │ 44.28 │ 0.22 │ 0.01 │

│ 3 │ 46.00 │ 47.81 │ -1.81 │ -0.04 │

│ 4 │ 56.50 │ 55.46 │ 1.04 │ 0.02 │

│ 5 │ 48.50 │ 45.15 │ 3.35 │ 0.07 │

│ 6 │ 47.20 │ 43.81 │ 3.39 │ 0.07 │

│ 7 │ 51.00 │ 48.27 │ 2.73 │ 0.05 │

│ 8 │ 37.00 │ 46.46 │ -9.46 │ -0.26 │

│ 9 │ 54.00 │ 54.57 │ -0.57 │ -0.01 │

│ 10 │ 42.20 │ 45.74 │ -3.54 │ -0.08 │

│ 11 │ 45.00 │ 53.34 │ -8.34 │ -0.19 │

│ 12 │ 64.50 │ 60.45 │ 4.05 │ 0.06 │

│ 13 │ 60.60 │ 60.47 │ 0.13 │ 0.00 │

│ 14 │ 52.00 │ 56.14 │ -4.14 │ -0.08 │

│ 15 │ 53.30 │ 53.89 │ -0.59 │ -0.01 │

│ 16 │ 57.80 │ 57.35 │ 0.45 │ 0.01 │

│ 17 │ 53.00 │ 52.31 │ 0.69 │ 0.01 │

│ 18 │ 61.50 │ 50.44 │ 11.06 │ 0.18 │

│ 19 │ 53.30 │ 52.63 │ 0.67 │ 0.01 │

│ 20 │ 52.00 │ 54.13 │ -2.13 │ -0.04 │

│ 21 │ 48.50 │ 50.59 │ -2.09 │ -0.04 │

│ 22 │ 52.30 │ 52.46 │ -0.16 │ -0.00 │

│ 23 │ 50.60 │ 50.03 │ 0.57 │ 0.01 │

│ 24 │ 51.00 │ 48.64 │ 2.36 │ 0.05 │

│ 25 │ 60.80 │ 63.60 │ -2.80 │ -0.05 │

└────┴──────────────┴───────────┴────────────┴───────────────┘

Характеристики остатков

Среднее значение..................... -0.000

Оценка дисперсии..................... 16.4

Оценка приведенной дисперсии........ 20.5

Средний модуль остатков.............. 2.850

Относительная ошибка аппроксимации... 0.057

Критерий Дарбина-Уотсона............. 1.845

Коэффициент детерминации............. 0.625

F - значение ( n1 = 5, n2 = 20)... 672

Гипотеза о значимости уравнения

не отвергается с вероятностью 0.950

\*\*\* Протокол множественной линейной регрессии \*\*\*

Зависимая переменная Y - y

Функция Y = +59.036+0.066\*x2+0.168\*x3-0.191\*x4

Оценки коэффициентов линейной регрессии

┌───┬──────────┬───────────┬───────────────┬───────────┬────────┬─────────┐

│ N │ Значение │ Дисперсия │ Средне- │ t - │ Нижняя │ Верхняя │

│ │ │ │ квадатическое │ значение │ оценка │ оценка │

│ │ │ │ отклонение │ │ │ │

├───┼──────────┼───────────┼───────────────┼───────────┼────────┼─────────┤

│ 1 │ 59.04 │ 15.07 │ 3.88 │ 15.21 │ 52.34 │ 65.74 │

│ 2 │ 0.07 │ 0.01 │ 0.08 │ 0.87 │ -0.07 │ 0.20 │

│ 3 │ 0.17 │ 0.01 │ 0.07 │ 2.28 │ 0.04 │ 0.30 │

│ 4 │ -0.19 │ 0.00 │ 0.05 │ -3.67 │ -0.28 │ -0.10 │

└───┴──────────┴───────────┴───────────────┴───────────┴────────┴─────────┘

Кpитические значения t-pаспpеделения

пpи 21 степенях свободы

веpоятность t-значение

0.900 1.326

0.950 1.726

* 1. 2.521

Оценки коэффициентов интерпретации линейной регрессии

╔════╤════════╤═════════╤═════════╗

║ N │Коэффиц.│Вета- │Дельта- ║

║ │эластичн│коэффиц. │коэффиц. ║

╠════╪════════╪═════════╪═════════╣

║1 │ +0.037│ +0.174│ +0.148║

║2 │ +0.073│ +0.458│ +0.444║

║3 │ -0.240│ -0.493│ +0.407║

╚════╧════════╧═════════╧═════════╝

Таблица остатков

┌────┬──────────────┬───────────┬────────────┬───────────────┐

│ N │ Эмпирическое │ Расчетное │ Ошибка │ Ошибка │

│ │ значение │ значение │ абсолютная │ относительная │

├────┼──────────────┼───────────┼────────────┼───────────────┤

│ 1 │ 63.00 │ 57.99 │ 5.01 │ 0.08 │

│ 2 │ 44.50 │ 44.04 │ 0.46 │ 0.01 │

│ 3 │ 46.00 │ 47.97 │ -1.97 │ -0.04 │

│ 4 │ 56.50 │ 55.68 │ 0.82 │ 0.01 │

│ 5 │ 48.50 │ 45.05 │ 3.45 │ 0.07 │

│ 6 │ 47.20 │ 43.55 │ 3.65 │ 0.08 │

│ 7 │ 51.00 │ 48.42 │ 2.58 │ 0.05 │

│ 8 │ 37.00 │ 46.53 │ -9.53 │ -0.26 │

│ 9 │ 54.00 │ 54.40 │ -0.40 │ -0.01 │

│ 10 │ 42.20 │ 45.73 │ -3.53 │ -0.08 │

│ 11 │ 45.00 │ 53.43 │ -8.43 │ -0.19 │

│ 12 │ 64.50 │ 60.78 │ 3.72 │ 0.06 │

│ 13 │ 60.60 │ 59.30 │ 1.30 │ 0.02 │

│ 14 │ 52.00 │ 55.72 │ -3.72 │ -0.07 │

│ 15 │ 53.30 │ 54.12 │ -0.82 │ -0.02 │

│ 16 │ 57.80 │ 57.55 │ 0.25 │ 0.00 │

│ 17 │ 53.00 │ 52.38 │ 0.62 │ 0.01 │

│ 18 │ 61.50 │ 50.53 │ 10.97 │ 0.18 │

│ 19 │ 53.30 │ 52.23 │ 1.07 │ 0.02 │

│ 20 │ 52.00 │ 54.43 │ -2.43 │ -0.05 │

│ 21 │ 48.50 │ 50.71 │ -2.21 │ -0.05 │

│ 22 │ 52.30 │ 52.56 │ -0.26 │ -0.01 │

│ 23 │ 50.60 │ 50.25 │ 0.35 │ 0.01 │

│ 24 │ 51.00 │ 48.87 │ 2.13 │ 0.04 │

│ 25 │ 60.80 │ 63.86 │ -3.06 │ -0.05 │

└────┴──────────────┴───────────┴────────────┴───────────────┘

Характеристики остатков

Среднее значение..................... -0.000

Оценка дисперсии..................... 16.5

Оценка приведенной дисперсии........ 19.7

Средний модуль остатков.............. 2.910

Относительная ошибка аппроксимации... 0.058

Критерий Дарбина-Уотсона............. 1.807

Коэффициент детерминации............. 0.622

F - значение ( n1 = 4, n2 = 21)... 876

Гипотеза о значимости уравнения

не отвергается с вероятностью 0.950

33333333333333333333 ОТЧЕТ-ИТОГ 33333333333333333333

\*\*\* Протокол множественной линейной регрессии \*\*\*

Зависимая переменная Y - y

Функция Y = +59.951+0.215\*x3-0.192\*x4

Оценки коэффициентов линейной регрессии

┌───┬──────────┬───────────┬───────────────┬───────────┬────────┬─────────┐

│ N │ Значение │ Дисперсия │ Средне- │ t - │ Нижняя │ Верхняя │

│ │ │ │ квадатическое │ значение │ оценка │ оценка │

│ │ │ │ отклонение │ │ │ │

├───┼──────────┼───────────┼───────────────┼───────────┼────────┼─────────┤

│ 1 │ 59.95 │ 13.80 │ 3.71 │ 16.14 │ 53.55 │ 66.35 │

│ 2 │ 0.22 │ 0.00 │ 0.05 │ 4.40 │ 0.13 │ 0.30 │

│ 3 │ -0.19 │ 0.00 │ 0.05 │ -3.71 │ -0.28 │ -0.10 │

└───┴──────────┴───────────┴───────────────┴───────────┴────────┴─────────┘

Кpитические значения t-pаспpеделения

пpи 22 степенях свободы

веpоятность t-значение

0.900 1.324

0.950 1.722

0.990 2.511

Оценки коэффициентов интерпретации линейной регрессии

╔════╤════════╤═════════╤═════════╗

║ N │Коэффиц.│Вета- │Дельта- ║

║ │эластичн│коэффиц. │коэффиц. ║

╠════╪════════╪═════════╪═════════╣

║1 │ +0.094│ +0.587│ +0.582║

║2 │ -0.241│ -0.495│ +0.418║

╚════╧════════╧═════════╧═════════╝

Таблица остатков

┌────┬──────────────┬───────────┬────────────┬───────────────┐

│ N │ Эмпирическое │ Расчетное │ Ошибка │ Ошибка │

│ │ значение │ значение │ абсолютная │ относительная │

├────┼──────────────┼───────────┼────────────┼───────────────┤

│ 1 │ 63.00 │ 56.40 │ 6.60 │ 0.10 │

│ 2 │ 44.50 │ 44.10 │ 0.40 │ 0.01 │

│ 3 │ 46.00 │ 47.16 │ -1.16 │ -0.03 │

│ 4 │ 56.50 │ 56.30 │ 0.20 │ 0.00 │

│ 5 │ 48.50 │ 45.94 │ 2.56 │ 0.05 │

│ 6 │ 47.20 │ 44.22 │ 2.98 │ 0.06 │

│ 7 │ 51.00 │ 47.97 │ 3.03 │ 0.06 │

│ 8 │ 37.00 │ 46.88 │ -9.88 │ -0.27 │

│ 9 │ 54.00 │ 54.68 │ -0.68 │ -0.01 │

│ 10 │ 42.20 │ 45.53 │ -3.33 │ -0.08 │

│ 11 │ 45.00 │ 54.55 │ -9.55 │ -0.21 │

│ 12 │ 64.50 │ 60.18 │ 4.32 │ 0.07 │

│ 13 │ 60.60 │ 59.44 │ 1.16 │ 0.02 │

│ 14 │ 52.00 │ 54.60 │ -2.60 │ -0.05 │

│ 15 │ 53.30 │ 53.21 │ 0.09 │ 0.00 │

│ 16 │ 57.80 │ 57.93 │ -0.13 │ -0.00 │

│ 17 │ 53.00 │ 51.65 │ 1.35 │ 0.03 │

│ 18 │ 61.50 │ 51.91 │ 9.59 │ 0.16 │

│ 19 │ 53.30 │ 52.71 │ 0.59 │ 0.01 │

│ 20 │ 52.00 │ 55.28 │ -3.28 │ -0.06 │

│ 21 │ 48.50 │ 50.61 │ -2.11 │ -0.04 │

│ 22 │ 52.30 │ 51.59 │ 0.71 │ 0.01 │

│ 23 │ 50.60 │ 50.86 │ -0.26 │ -0.01 │

│ 24 │ 51.00 │ 48.09 │ 2.91 │ 0.06 │

│ 25 │ 60.80 │ 64.35 │ -3.55 │ -0.06 │

└────┴──────────────┴───────────┴────────────┴───────────────┘

Характеристики остатков

Среднее значение..................... 0.000

Оценка дисперсии..................... 17.1

Оценка приведенной дисперсии........ 19.4

Средний модуль остатков.............. 2.920

Относительная ошибка аппроксимации... 0.058

Критерий Дарбина-Уотсона............. 1.864

Коэффициент детерминации............. 0.609

F - значение ( n1 = 3, n2 = 22)... 1.18e+03

Гипотеза о значимости уравнения

не отвергается с вероятностью 0.950

# ***Использованная литература:***

1. Френкель А.А., Адамова Е.В. Корреляционный и регрессионный анализ в экономических приложениях: Учебное пособие / МЕСИ – М:, 1987 г.
2. Мхитарян В.С., Трошин Л.И., Адамова Е.В., Шевченко К.К., Бамбаева Н.Я. Теория вероятностей и математическая статистика / Московский международный институт эконометрики, информатики, финансов и права. – М., 2002 г.
3. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – М., ЮНИТИ-ДАНА, 2001 г.
4. Колемаев В.А. и др. Теория вероятностей и математическая статистика. Учеб. пособие для экон. спец. вузов. – М., Высшая школа, 1991 г.
5. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие для втузов. Изд. 5-е, переработанное и дополненное. М., Высшая школа, 1977 г.