# РЕГРЕССИОННЫЙ АНАЛИЗ

## 1.Общая схема регрессионного анализа

1. Выбор аппроксимирующей функции

В тех случаях, когда математическая модель процесса неизвестна, для аппроксимации имеющихся данных используются полиномиальные зависимости. Для одномерного случая такой полином имеет вид:

, (4.1)

где

*y* – выходной параметр;

*x* – фактор (входной параметр);

*а* – коэффициенты уравнения;

*n* – степень полинома.

Например:

полином первой степени имеет вид:  (4.2)

полином второй степени:  (4.3)

полином третьей степени:  (4.4)

и т.д.

Полиномы более высоких степеней использовать не рекомендуется, поскольку даже если при последующем анализе окажутся статистически значимыми коэффициенты при высоких степенях, то их физический смысл будет очень трудно объяснить.

2. Расчет коэффициентов полинома

Вычисление коэффициентов удобнее всего производить используя аппарат матричной алгебры.

Основное уравнение для вычисления коэффициентов в матричном виде выглядит следующим образом:

, (4.5)

где – матрица входных параметров;

- транспонированная матрица входных параметров;

- обратная матрица произведения ;

- вектор-столбец значений выходного параметра.

3.Проверка уравнения на адекватность

Суть проверки заключается в сравнении дисперсии выходного параметра, обусловленной применяемым уравнением с дисперсией ошибки измерений.

В случае технических экспериментов ошибка измерений оценивается по результатам повторных измерений при одних и тех же значениях входных параметров.

Для экономических данных повторные измерения как правило невозможны. Поэтому в таких случаях ошибка измерений оценивается как разность между общей дисперсией выходного параметра и дисперсией, которая обусловлена применением уравнения, т.е.:

 (4.6)

где – сумма квадратов отклонений значений выходного параметра от его среднего значения;

- сумма квадратов отклонений значений выходного параметра от рассчитанных значений;

- сумма квадратов отклонений, вызванная применением уравнения.

Далее находятся средние дисперсии для ошибки измерений и уравнения. Вычисления производятся по следующим формулам:

для ошибки: ; (4.7)

для уравнения: , (4.8)

где *f –* число степеней свободы. Этот параметр вычисляется по формулам:

для уравнения: *fуравнения=k+1*; (4.9)

для ошибки: *fошибки=N-k-1*, (4.10)

где *N* - общее число измерений;

*k* - количество коэффициентов в уравнении.

Для проверки адекватности уравнения вычисляется критерий Фишера:

, (4.11)

Который затем сравнивается с критическим значением критерия при выбранном уровне значимости (стандартным является уровень значимости, равный 0,05) и числе степеней свободы для числителя и знаменателя.

Если рассчитанное значение критерия больше критического, то уравнения признается адекватным. В противном случае уравнение является не адекватным.

4. Проверка коэффициентов на значимость

Проверка производится следующим образом.

Сначала вычисляется дисперсия каждого коэффициента уравнения:

, (4.12)

где  - значение *i*–ого диагонального элемента матрицы .

Далее вычисляется стандартное отклонение каждого коэффициента:

 (4.13)

Затем для каждого коэффициента вычисляется доверительный интервал:

, (4.14)

где  *t* – критерий Стьюдента при выбранном уровне значимости и числе степеней, равном *fошибки*.

Если значение коэффициента больше его доверительного интервала, то данный коэффициент признается статистически значимым. В противном случае коэффициент является не значимым.

Далее незначимые коэффициенты удаляются из исходного уравнения и для упрощенного таким образом уравнения производятся повторные расчеты согласно этапам 1 – 4.

5. Прогнозирование и оценка ошибки прогноза

Вычисление выходного параметра в предсказываемой точке факторного пространства производится по формуле:

, (4.15)

где  - значение выходного параметра в точке предсказания;

 - транспонированный вектор входного параметра в точке предсказания;

*А* - вектор коэффициентов.

Определение дисперсии выходного параметра в точке предсказания производится по формуле:

 (4.16)

Доверительный интервал выходного параметра в точке предсказания:

, (4.17)

где *t* - критерий Стьюдента при выбранном уровне значимости и числе степеней свободы, равном *N-k-1*.

## 2. Пример

Используем для обработки данные о наличие квартирных телефонных аппаратов телефонной сети общего пользования или имеющих на нее выход на 1000 чел. сельского населения (на конец года; шт.).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1995 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 |
| 41,0 | 47,2 | 49,1 | 51,8 | 56,2 | 59,7 | 63,3 | 67,8 | 72,7 |

Количество телефонных аппаратов будет являться выходным параметром. Время будет выступать в качестве входного параметра.

Для расчетов будем использовать данные за 1995 – 2003 г. Данные за 2004 г. будут использованы для сопоставления с результатами прогноза.

**Выбор уравнение регрессии**

В качестве уравнения регрессии выберем полином третьей степени:

 (4.18)

**Формирование матрицы переменных**

Для выполнения расчетов разместим данные следующим образом:

|  |  |  |  |
| --- | --- | --- | --- |
|  | B | C | D |
| 4 |  |  |  |
| 5 |  | Y | X |
| 6 |  | 41 | 1995 |
| 7 |  | 47,2 | 1997 |
| 8 |  | 49,1 | 1998 |
| 9 |  | 51,8 | 1999 |
| 10 |  | 56,2 | 2000 |
| 11 |  | 59,7 | 2001 |
| 12 |  | 63,3 | 2002 |
| 13 |  | 67,8 | 2003 |

До начала расчетов входную переменную необходимо нормировать и центрировать. Нормирование и центрирование производится по формуле:

, (4.19)

где - кодированное значение входного параметра;

- натуральное значение входного параметра;

- максимальное и минимальное значение входного параметра;

- центральное значение входного параметра.

Центральное значение параметра определяется по формуле:

 (4.20)

Для реализации расчетов по формулам (4.19)-(4.20):

в ячейку D14 вводится формула =МИН(D6:D13);

в ячейку D15 вводится формула =МАКС(D6:D13);

в ячейку D16 вводится формула =(D14+D15)/2;

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | B | C | D | E | F | G | H |
| 3 |  |  |  |  |  |  |  |
| 4 |  |  | Натуральные значения | Кодированные значения | | | |
| 5 |  | Y | X | x0 | x1 | x2 | x3 |
| 6 |  | 41 | 1995 | 1 | -1 | 1 | -1 |
| 7 |  | 47,2 | 1997 | 1 | -0,5 | 0,25 | -0,125 |
| 8 |  | 49,1 | 1998 | 1 | -0,25 | 0,0625 | -0,016 |
| 9 |  | 51,8 | 1999 | 1 | 0 | 0 | 0 |
| 10 |  | 56,2 | 2000 | 1 | 0,25 | 0,0625 | 0,0156 |
| 11 |  | 59,7 | 2001 | 1 | 0,5 | 0,25 | 0,125 |
| 12 |  | 63,3 | 2002 | 1 | 0,75 | 0,5625 | 0,4219 |
| 13 |  | 67,8 | 2003 | 1 | 1 | 1 | 1 |
| 14 |  | Мин= | 1995 |  |  |  |  |
| 15 |  | Макс= | 2003 |  |  |  |  |
| 16 |  | Центр= | 1999 |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

- ячейки E6:E13 соответствуют фиктивной переменной - коэффициенту а0 уравнения (18) и заполняются единицами:

- в ячейку F6 вводится формула =(D6-$D$16)/($D$15-$D$14)\*2;

- в ячейку G6 вводится формула =F6^2;

- в ячейку H6 вводится формула =F6^3.

Все введенные формулы копируются до 13 строки.

В результате пересчета получены кодированные значения входных переменных с диапазоном изменения -1..+1.

Для удобства дальнейшей работы диапазону C6:C13 присвоим имя Y, а диапазону E6:H13 имя X.

**Расчет коэффициентов уравнения**

Расчет коэффициентов производится согласно уравнению (4.5).

Выполним поэтапное вычисление по указанной формуле.

1. Получим транспонированную матрицу X.

Для этого перейдем на ячейку E19 и вызовем функцию ТРАНСП. В окне параметров функции в качестве «массив» указать X,

При этом в ячейке E19 появится сообщение «#ЗНАЧ!». Выделим диапазон L19:K22, затем нажимаем F2 и выполним тройное нажатие Ctrl+Shift+Enter.

\*Размер выделяемого диапазона (количество строк и столбцов) должен соответствовать размерности матрицы X - т.е. количество выделенных строк должно быть равно количеству коэффициентов в уравнении регрессии, а количество столбцов количеству данных. Чтобы не ошибиться при выделении рекомендуется заранее пронумеровать нужные строки и столбцы.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | D | E | F | G | H | I | J | K | L |  |
| 18 |  |  |  |  |  |  |  |  |  |  |
| 19 |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |  |
| 20 | Xt= | -1 | -0,5 | -0,25 | 0 | 0,25 | 0,5 | 0,75 | 1 |  |
| 21 |  | 1 | 0,25 | 0,0625 | 0 | 0,0625 | 0,25 | 0,5625 | 1 |  |
| 22 |  | -1 | -0,125 | -0,01563 | 0 | 0,015625 | 0,125 | 0,421875 | 1 |  |
|  |  |  |  |  |  |  |  |  |  |  |

Присвоим этому диапазону имя Xt.

Выполним умножение матриц  и 

Для этого перейдем на ячейку E25 и вызовем функцию МУМНОЖ. В окне параметров функции в качестве «массив 1» указать Xt, а в качестве «массив 2» указать X.

При этом в ячейке E25 появится число 8. Выделим диапазон E25:H28\*, затем нажимаем F2 и выполним тройное нажатие Ctrl+Shift+Enter.

\*Размер выделяемого диапазона (количество строк и столбцов) должен соответствовать количеству коэффициентов в уравнении регрессии.

В результате в диапазоне E25:H28 появится результат перемножения:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | D | E | F | G | H | I |
| 24 |  |  |  |  |  |  |  |
| 25 |  |  | 8 | 0,75 | 3,1875 | 0,421875 |  |
| 26 |  |  | 0,75 | 3,1875 | 0,421875 | 2,449219 |  |
| 27 |  | XtX= | 3,1875 | 0,421875 | 2,449219 | 0,237305 |  |
| 28 |  |  | 0,421875 | 2,449219 | 0,237305 | 2,209717 |  |
| 29 |  |  |  |  |  |  |  |

Присвоим этому диапазону имя XtX.

Выполним обращение этой матрицы.

Для этого перейдем на ячейку E31 и вызовем функцию МОБР. В качестве аргумента укажем XtX.

При этом в ячейке E31 появится число 0,2609.

Выделим диапазон E31:H34, затем нажимаем F2 и выполним тройное нажатие Ctrl+Shift+Enter.

В указанных ячейках должен появиться результат обращения матрицы:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | C | D | E | F | G | H | I |
| 30 |  |  |  |  |  |  |  |
| 31 |  |  | 0,2609 | -0,04314 | -0,33538 | 0,034024 |  |
| 32 |  | XtXobr= | -0,04314 | 2,169717 | -0,08628 | -2,38738 |  |
| 33 |  |  | -0,33538 | -0,08628 | 0,853042 | 0,068049 |  |
| 34 |  |  | 0,034024 | -2,38738 | 0,068049 | 3,084884 |  |
| 35 |  |  |  |  |  |  |  |

Присвоим этому диапазону имя XtXobr.

Выполним умножение матриц Xt и Y.

Для этого переходим на ячейку E37 и вызовем функцию МУМНОЖ. В окне параметров функции в качестве «массив 1» указать Xt, а в качестве «массив 2» указать Y.

При этом в ячейке E37 появится число 436,1.

Выделим диапазон E37:E40, затем нажимаем F2 и выполним тройное нажатие Ctrl+Shift+Enter.

\*Количество выделенных строк соответствует количеству коэффициентов в уравнении регрессии.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | C | D | E | F |
| 36 |  |  |  |  |
| 37 |  |  | 436,1 |  |
| 38 |  | XtY= | 82,3 |  |
| 39 |  |  | 177,7125 |  |
| 40 |  |  | 55,17813 |  |
| 41 |  |  |  |  |

Диапазону E37:E40 присвоим имя XtY.

И наконец последнее вычисление по формуле (4.5), т.е. перемножим матрицы XtXobr и XtY.

Для этого переходим на ячейку E43 и вызовем функцию МУМНОЖ. В окне параметров функции в качестве «массив 1» указать XtXobr, а в качестве «массив 2» указать XtY.

При этом в ячейке E43 появится число 52,49476.

Выделим диапазон E43:E46\*, затем нажимаем F2 и выполним тройное нажатие Ctrl+Shift+Enter.

\*Количество выделенных строк соответствует количеству коэффициентов в аппроксимирующем полиноме.

Должен получиться следующий результат:

|  |  |  |  |
| --- | --- | --- | --- |
|  | C | D | E |
| 42 |  |  |  |
| 43 |  |  | 52,49476 |
| 44 |  | A= | 12,69151 |
| 45 |  |  | 1,989519 |
| 46 |  |  | 0,66773 |
| 47 |  |  |  |

Диапазону E43:E46 присвоим имя A.

Таким образом у нас получилось уравнение:

Y=52,49+12,69x+1,99\*x^2+0,67\*x^3 (4.21)

**Проверка адекватности уравнения**

Проверка адекватности производится на основе формул (4.6) - (4.11).

Реализация расчетов в Excel выглядит следующим образом:

1. Сначала вычисляются компоненты формулы (4.6) (эти компоненты называются также остаточными суммами квадратов).

Вычисление общей остаточной суммы квадратов производится по формуле:

, (4.22)

где n - число измерений;

- *i*-ое значение выходного параметра;

 - среднее значение выходного параметра.

Остаточная сумма квадратов, обусловленная ошибкой измерений после применения модели вычисляется по внешне аналогичной формуле:

, (4.23)

где  - значение выходного параметра, рассчитанное с помощью уравнения регрессии в *i*-ой точке.

Остаточная сумма квадратов, обусловленная моделью вычисляется как разность первых двух уже определенных компонентов.

Для вычислений по формуле (4.23) предварительно рассчитаем значения выходного параметра по полученному уравнению:

в ячейку I6 вводится формула:

=МУМНОЖ(E6:H6;A),

которая копируется до 13 строки столбца I.

Далее выполним расчеты в следующих ячейках:

В ячейках C49, D49 указываются степени используемого уравнения.

В ячейку C50 вводится формула:

=ДИСП(Y)\*7,

а в ячейку D50 вводится формула:

=СУММКВРАЗН(Y;I6:I13).

В ячейку D51

=C50-D50.

В ячейке D52 рассчитывается число степеней свободы для ошибки. Расчет производится по формуле:

f=N-1-k=8-1-4=3,

где N - общее число измерений;

k - количество коэффициентов в уравнении.

В ячейку D53 рассчитывается число степеней свободы для уравнения =4.

В ячейке D54 рассчитывается дисперсия ошибки:

=D50/D52.

В ячейке D55 рассчитывается дисперсия уравнения:

=D51/D53.

В ячейке D56 рассчитывается критерий Фишера:

=D55/D54.

Рассчитанное значение критерия сравнивается с критическим при заданном уровне значимости (0,05) и числе степеней свободы для числителя, равном k, знаменателя, равном N-1-k.

Критическое значение можно определить с помощью функции Excel =FРАСПОБР(0,05;4;3). В данном случае оно будет равно 9,11 (ячейка D57).

Результаты вычислений представлены в следующей таблице:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | A | B | C | D |
| 48 |  |  |  |  |
| 49 |  | Степень полинома | 0 | 3 |
| 50 |  | Остаточная сумма квадратов | 556,25 | 1,2676 |
| 51 |  | Сумма квадратов, приходящаяся на уравнение |  | 554,98 |
| 52 |  | Число степеней свободы ошибки |  | 3 |
| 53 |  | Число степеней свободы уравнения |  | 4 |
| 54 |  | Дисперсия ошибки |  | 0,4225 |
| 55 |  | Дисперсия уравнения |  | 138,75 |
| 56 |  | Критерий Фишера |  | 328,36 |
| 57 |  | Критическое значение Фишера |  | 9,1172 |
|  |  |  |  |  |

Рассчитанное значение критерия Фишера больше критического. Поэтому делается вывод о том, что полученное уравнение является адекватным.

**Проверка коэффициентов модели на статистическую значимость**

Проверка значимости коэффициентов производится на основе вычислений по формулам (4.12)-(4.14).

Сами вычисления удобно представить в виде таблицы:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | D | E | F | G | H | I |
| 63 |  | Значение коэффициентов | Значение диагональных элементов матрицы | Дисперсии коэффициентов | Стандартное отклонение коэффициентов | Доверительные интервалы коэффициентов |
| 64 |  | 52,49475 | 0,260879471 | 0,11023064 | 0,332010002 | 0,785078902 |
| 65 |  | 12,6915098 | 2,169717222 | 0,91678092 | 0,95748677 | 2,264096437 |
| 66 |  | 1,98951923 | 0,853041695 | 0,36043976 | 0,600366351 | 1,419640834 |
| 67 |  | 0,66772931 | 3,084883927 | 1,30347065 | 1,14169639 | 2,699682971 |
| 68 |  |  |  |  |  |  |

В ячейки E64:E67 копируются значения коэффициентов.

В ячейки F64:F67 копируются диагональные элементы матрицы вводится ;

в ячейку G64 вводится формула (4.12), т.е. =F64\*$D$54;

в ячейку H64 вводится формула (4.13), т.е. =G64^0,5;

в ячейку I64 вводится формула (4.14), т.е. =H64\*$I$61

При этом в ячейку I61 введена формула, которая вычисляет критическое значение критерия Стьюдента:

=СТЬЮДЕНТ.ОБР.2Х(0,05;7).

Все указанные формулы копируются до 67 строки.

Определение статистической значимости коэффициентов определяется путем сравнения значений каждого коэффициента с соответствующим доверительным интервалом:

- если значение коэффициента по модулю больше величины доверительного интервала, то данный коэффициент признается статистически значимым. В противном случае коэффициент считается незначимым.

Прямое сравнение столбцов E и I в вышеприведенной таблице показывает, в уравнении (4.21) статистически значимыми являются коэффициенты a0, a1 и а2.

Основным следствием этого результата является следующее:

- в качестве уравнения регрессии необходимо взять более простое уравнение:

y=a0+a1\*x+а2\*x^2 (4.24)

и произвести перерасчет его коэффициентов по описанной выше схеме.

**Результаты пересчета по уравнению (4.24)**

Пересчет удобнее произвести на новом листе Excel.

Далее без подробных комментариев приводятся только основные результаты такого пересчета.

Матрица X:

|  |  |  |
| --- | --- | --- |
| x0 | x | X^2 |
| 1 | -1 | 1 |
| 1 | -0,5 | 0,25 |
| 1 | -0,25 | 0,0625 |
| 1 | 0 | 0 |
| 1 | 0,25 | 0,0625 |
| 1 | 0,5 | 0,25 |
| 1 | 0,75 | 0,5625 |
| 1 | 1 | 1 |

Транспонированная матрица X:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Xt= | -1 | -0,5 | -0,25 | 0 | 0,25 | 0,5 | 0,75 | 1 |
|  | 1 | 0,25 | 0,0625 | 0 | 0,0625 | 0,25 | 0,5625 | 1 |

Матрица XtX

|  |  |  |  |
| --- | --- | --- | --- |
| XtX= | 8 | 0,75 | 3,1875 |
|  | 0,75 | 3,1875 | 0,421875 |
|  | 3,1875 | 0,421875 | 2,449219 |

Обратная матрица (XtX)-1:

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0,260504 | -0,01681 | -0,33613 |
| XtXobr= | -0,01681 | 0,322129 | -0,03361 |
|  | -0,33613 | -0,03361 | 0,851541 |

Матрица XtY:

|  |  |
| --- | --- |
|  | 436,1 |
| XtY= | 82,3 |
|  | 177,7125 |

Коэффициенты (матрица А):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | B | C | D | E |
|  |  |  |  |  |
| 34 |  |  | 52,48739 |  |
| 35 |  | A= | 13,20826 |  |
| 36 |  |  | 1,97479 |  |

Результат проверки уравнения на адекватность:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | J | K | L | M |
|  |  |  |  |  |
| 42 |  | Степень полинома | 0 | 3 |
| 43 |  | Остаточная сумма квадратов | 556,2487 | 1,412136 |
| 44 |  | Сумма квадратов, приходящаяся на уравнение |  | 554,8366 |
| 45 |  | Число степеней свободы ошибки |  | 4 |
| 46 |  | Число степеней свободы уравнения |  | 3 |
| 47 |  | Дисперсия ошибки |  | 0,353034 |
| 48 |  | Дисперсия уравнения |  | 184,9455 |
| 49 |  | Критерий Фишера |  | 523,8746 |
| 50 |  | Критическое значение Фишера |  | 6,591382 |

Уравнение адекватно.

D52 вводится формула для определения критического значения коэффициента Стьюдента =СТЬЮДРАСПОБР(0,05;7).

Проверка коэффициентов уравнения на значимость:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Значение коэффициентов | Значение диагональных элементов матрицы | Дисперсии коэффициентов | Стандартное отклонение коэффициентов | Доверительные интервалы коэффициентов |
|  |  |  |  |  |
| 52,48739 | 0,260504 | 0,09196683 | 0,30326034 | 0,717096 |
| 13,20826 | 0,322129 | 0,11372243 | 0,33722756 | 0,797417 |
| 1,97479 | 0,851541 | 0,30062276 | 0,54829076 | 1,296502 |
|  |  |  |  |  |
|  |  |  |  |  |

Все коэффициенты статистически значимы.

Таким образом уравнение, описывающее наши данные имеет вид:

+1,97\*x^2 (4.25)

**Прогнозирование и оценка ошибки прогноза**

Согласно постановке задачи нам необходимо спрогнозировать количество телефонных аппаратов в 2004 году.

Для этого расширим матрицу исходных данных с учетом лет для прогноза и в столбце G рассчитаем значения Y по полученному уравнению.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **C** | **D** | **E** | **F** | **G** |
|  |  |  |  |  |  |  |
| 4 | Yreal | Годы | x0 | x | X^2 | Yrasch |
| 5 | 41 | 1995 | 1 | -1 | 1 | 41,2539216 |
| 6 | 47,2 | 1997 | 1 | -0,5 | 0,25 | 46,3769608 |
| 7 | 49,1 | 1998 | 1 | -0,25 | 0,0625 | 49,3087535 |
| 8 | 51,8 | 1999 | 1 | 0 | 0 | 52,487395 |
| 9 | 56,2 | 2000 | 1 | 0,25 | 0,0625 | 55,9128852 |
| 10 | 59,7 | 2001 | 1 | 0,5 | 0,25 | 59,5852241 |
| 11 | 63,3 | 2002 | 1 | 0,75 | 0,5625 | 63,5044118 |
| 12 | 67,8 | 2003 | 1 | 1 | 1 | 67,6704482 |
| 13 | 72,7 | 2004 | 1 | 1,25 | 1,5625 | 72,08333 |

Для этого в ячейку G5 вводится формула:

=МУМНОЖ(E5:G5;$D$34:$D$36),

которая затем скопирована до 13 строки.

Здесь $D$34:$D$36 – адрес коэффициентов нового (квадратичного) полинома.

Ошибка прогноза рассчитывается по формулам (4.16) - (4.17).





Согласно формуле (4.16) необходимо сначала рассчитать прямой и транспонированный вектор входной переменной в точке прогноза. Транспонированный вектор уже известен – в ячейках D13F13. Прямой вектор разместим S2:S4:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | R | S | T |
| 2 |  | Xp= | 1 |  |
| 3 |  | 1,25 |  |
| 4 |  | 1,5625 |  |
| 5 |  | Xobr\*Xp= | -0,28571 |  |
| 6 |  | 0,333333 |  |
| 7 |  | 0,952381 |  |
| 8 |  | Xtp\*Xobr\*Xp= | 1,619048 |  |
| 9 |  | S2y= | 0,571579 |  |
| 10 |  | Sy=(S2y)^0,5= | 0,756028 |  |
| 11 |  | dy=Sy\*t | 1,787723 |  |
| 12 |  | Ymin= | 70,29561 |  |
| 13 |  | Ymax= | 73,87106 |  |
| 14 |  | Yreal= | 72,7 |  |
| 15 |  |  |  |  |

Затем перемножаем транспонированную матрицу входных переменных на обратную матрицу. Для этого в C5 вводится формула:

=МУМНОЖ($D$26:$F$28;S2:S4),

где ($D$26:$F$28 – адрес обратной матрицы.

- выделяем S5:S7, затем F2 и Ctrl+Shift+Enter.

В ячейку S8 вводим формулу: =МУМНОЖ(E13:G13;S5:S7).

В ячейку S9 вводим формулу: =S8\*$D$45 (где $D$45 адрес ячейки, в которой находится вычисленная дисперсия ошибки).

В ячейку S10 вводим формулу: =S9^0,5.

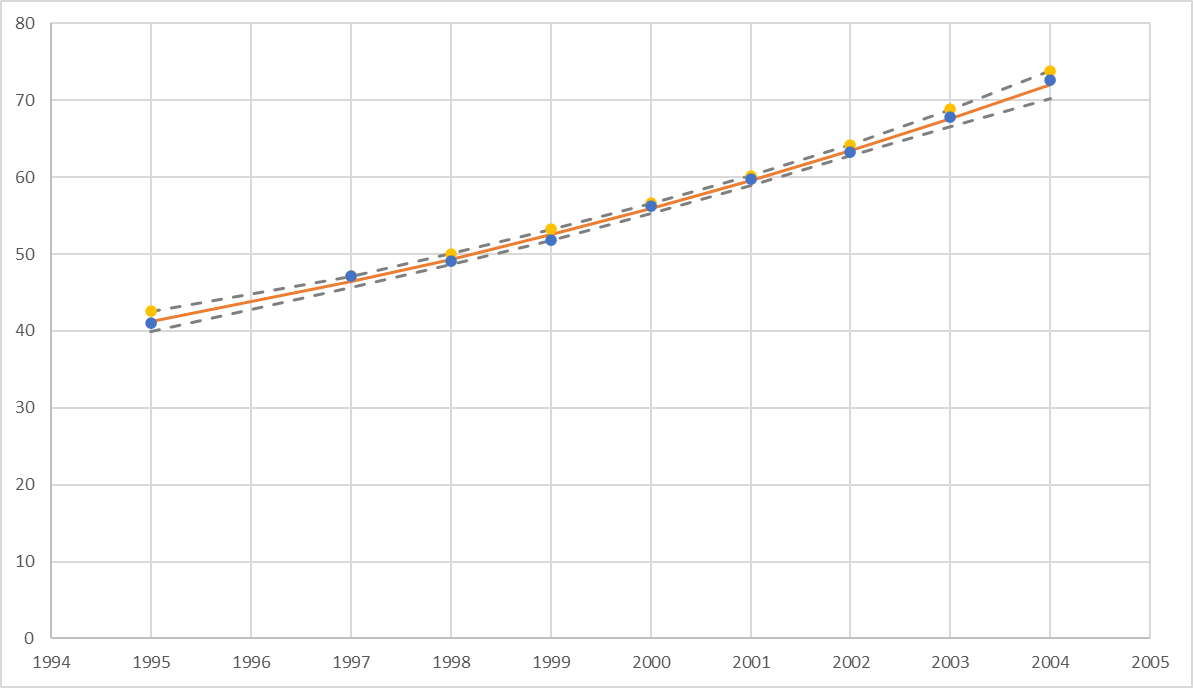
В ячейку S11 вводим формулу: =S10\*$K$21 (где $K$21 адрес ячейки с критерием Стьюдента).

В ячейках S12 и S13 рассчитываются нижняя и верхняя границы прогноза.

При этом в ячейку S12 введена формула: =H13-S11, а в S13 формула =H13\*S11.

Таким образом, предсказываемые значения показателя

- в 2004 равно 72,08 +/- 1,79 (реальное значение равно 72,7).



**3. Варианты заданий**

1. Данные для анализа находятся в файле «Данные по Чувашии.xls».
2. Во всех вариантах получить уравнение регрессии, адекватно описывающее представленные данные.
3. Для расчетов используются данные до 2017 г.
4. Данные за 2017 г. использовать для сопоставления с результатами прогноза.