5. Данные о состоянии уровня преступности в нашем городе за последние 15 месяцев представлены в таблице.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Месяц | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Уровень | 59 | 60 | 62 | 58 | 65 | 75 | 81 | 90 | 103 | 107 | 112 | 116 | 122 | 125 | 130 |

Определить оптимальный тренд и рассчитать точечный прогноз на последующие пять месяцев. Проверить модель на значимость.

Оптимальный тренд определим в Excel

Построим линейный тренд

Построим экспоненциальный тренд

Построим логарифмический тренд

Построим полиноминальный тренд

Построим степенной тренд

Как видим максимальные и практически одинаковые коэффициенты детерминации имеет линейный R^2 = 0,9626 и параболический R^2 = 0,9652 тренды.

Рассмотрим линейный тренд

Линейное уравнение тренда имеет вид y = bt + a

**Находим параметры уравнения методом наименьших квадратов**.

Система уравнений МНК:

an + b∑t = ∑y

a∑t + b∑t2 = ∑y\*t

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| t | y | t2 | y2 | t y |
| 1 | 59 | 1 | 3481 | 59 |
| 2 | 60 | 4 | 3600 | 120 |
| 3 | 62 | 9 | 3844 | 186 |
| 4 | 58 | 16 | 3364 | 232 |
| 5 | 65 | 25 | 4225 | 325 |
| 6 | 75 | 36 | 5625 | 450 |
| 7 | 81 | 49 | 6561 | 567 |
| 8 | 90 | 64 | 8100 | 720 |
| 9 | 103 | 81 | 10609 | 927 |
| 10 | 107 | 100 | 11449 | 1070 |
| 11 | 112 | 121 | 12544 | 1232 |
| 12 | 116 | 144 | 13456 | 1392 |
| 13 | 122 | 169 | 14884 | 1586 |
| 14 | 125 | 196 | 15625 | 1750 |
| 15 | 130 | 225 | 16900 | 1950 |
| 120 | 1365 | 1240 | 134267 | 12566 |
| Ср.знач. | 91 | 82.667 | 8951.133 | 837.733 |

Для наших данных система уравнений имеет вид:

15a + 120b = 1365

120a + 1240b = 12566

Из первого уравнения выражаем a и подставим во второе уравнение

Получаем a = 43.971, b = 5.879

Уравнение тренда:

y = 5.879 t + 43.971

Эмпирические коэффициенты тренда *a* и *b* являются лишь оценками теоретических коэффициентов βi, а само уравнение отражает лишь общую тенденцию в поведении рассматриваемых переменных.

Коэффициент тренда b = 5.879 показывает среднее изменение результативного показателя (в единицах измерения у) с изменением периода времени t на единицу его измерения. В данном примере с увеличением t на 1 единицу, y изменится в среднем на 5.879.

**Ошибка аппроксимации**.

Для оценки качества параметров уравнения построим расчетную таблицу.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| t | y | y(t) | (yi-ycp)2 | (yi-y(t))2 | (yi-y(t)) : yi |
| 1 | 59 | 49.85 | 1024 | 83.722 | 0.155 |
| 2 | 60 | 55.729 | 961 | 18.245 | 0.0712 |
| 3 | 62 | 61.607 | 841 | 0.154 | 0.00634 |
| 4 | 58 | 67.486 | 1089 | 89.979 | 0.164 |
| 5 | 65 | 73.364 | 676 | 69.961 | 0.129 |
| 6 | 75 | 79.243 | 256 | 18.002 | 0.0566 |
| 7 | 81 | 85.121 | 100 | 16.986 | 0.0509 |
| 8 | 90 | 91 | 1 | 1 | 0.0111 |
| 9 | 103 | 96.879 | 144 | 37.472 | 0.0594 |
| 10 | 107 | 102.757 | 256 | 18.002 | 0.0397 |
| 11 | 112 | 108.636 | 441 | 11.318 | 0.03 |
| 12 | 116 | 114.514 | 625 | 2.207 | 0.0128 |
| 13 | 122 | 120.393 | 961 | 2.583 | 0.0132 |
| 14 | 125 | 126.271 | 1156 | 1.617 | 0.0102 |
| 15 | 130 | 132.15 | 1521 | 4.622 | 0.0165 |
|   |   | 1365 | 10052 | 375.871 | 0.825 |

Оценим качество уравнения тренда с помощью средней относительной ошибки аппроксимации.

Ошибка аппроксимации в пределах 5%-7% свидетельствует о хорошем подборе уравнения тренда к исходным данным.

Поскольку ошибка меньше 7%, то данное уравнение можно использовать в качестве тренда.

Средние значения

Дисперсия

Среднеквадратическое отклонение

**Эмпирическое корреляционное отношение**.

Эмпирическое корреляционное отношение вычисляется для всех форм связи и служит для измерение тесноты зависимости. Изменяется в пределах [0;1].

=

где

В отличие от линейного коэффициента корреляции он характеризует тесноту нелинейной связи и не характеризует ее направление. Изменяется в пределах [0;1].

Связи между признаками могут быть слабыми и сильными (тесными). Их критерии оцениваются по шкале Чеддока:

0.1 < η < 0.3: слабая;

0.3 < η < 0.5: умеренная;

0.5 < η < 0.7: заметная;

0.7 < η < 0.9: высокая;

0.9 < η < 1: весьма высокая;

Полученная величина свидетельствует о том, что изменение временного периода t существенно влияет на y.

**Коэффициент детерминации**.

т.е. в 96.26% случаев t влияет на изменение y. Другими словами - точность подбора уравнения тренда - высокая.

**2. Анализ точности определения оценок параметров уравнения тренда**.

Дисперсия ошибки уравнения.

где m = 1 - количество влияющих факторов в модели тренда.

Стандартная ошибка уравнения.

**Интервальный прогноз**.

Определим среднеквадратическую ошибку прогнозируемого показателя.

Uy = yn+L ± K

где

L - период упреждения; уn+L - точечный прогноз по модели на (n + L)-й момент времени; n - количество наблюдений во временном ряду; Sy - стандартная ошибка прогнозируемого показателя; Tтабл - табличное значение критерия Стьюдента для уровня значимости α и для числа степеней свободы, равного *n-2*.

По таблице Стьюдента находим Tтабл

Tтабл (n-m-1;α/2) = (;) = 2.16

Точечный прогноз, t = 16: y(16) = 5.879\*16 + 43.971 = 138.03

138.03 - 13.22 = 124.81 ; 138.03 + 13.22 = 151.25

Интервальный прогноз:

t = 16: (124.81;151.25)

Точечный прогноз, t = 17: y(17) = 5.879\*17 + 43.971 = 143.91

143.91 - 13.52 = 130.39 ; 143.91 + 13.52 = 157.43

Интервальный прогноз:

t = 17: (130.39;157.43)

Точечный прогноз, t = 18: y(18) = 5.879\*18 + 43.971 = 149.79

149.79 - 13.86 = 135.93 ; 149.79 + 13.86 = 163.65

Интервальный прогноз:

t = 18: (135.93;163.65)

Точечный прогноз, t = 19: y(19) = 5.879\*19 + 43.971 = 155.66

155.66 - 14.22 = 141.44 ; 155.66 + 14.22 = 169.88

Интервальный прогноз:

t = 19: (141.44;169.88)

Точечный прогноз, t = 20: y(20) = 5.879\*20 + 43.971 = 161.54

161.54 - 14.6 = 146.94 ; 161.54 + 14.6 = 176.14

Интервальный прогноз:

t = 20: (146.94;176.14)

**F-статистика. Критерий Фишера.**

Коэффициент детерминации.

=

Находим из таблицы Fkp(1;13;0.05) = 4.67

где m - количество факторов в уравнении тренда (m=1).

Поскольку F > Fkp, то коэффициент детерминации (и в целом уравнение тренда) статистически значим.

**Выводы**

Изучена временная зависимость Y (уровня преступности) от времени t. На этапе спецификации был выбран линейный тренд. Оценены его параметры методом наименьших квадратов. Статистическая значимость уравнения проверена с помощью коэффициента детерминации и критерия Фишера. Установлено, что в исследуемой ситуации 96.26% общей вариабельности Y объясняется изменением временного параметра. Возможна интерпретация параметров модели - с каждым периодом времени t значение Y (уровень преступности) в среднем увеличивается на 5.879 ед. изм.

Рис. 1 Исходные данные, линейный тренд и прогноз