

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»

И. Н. Булгакова

**ЭКОНОМЕТРИКА:
ЛАБОРАТОРНЫЙ ПРАКТИКУМ**

Учебное пособие для вузов

Воронеж
Издательский дом ВГУ
2016

СОДЕРЖАНИЕ

| | |
|---|-----------|
| Лабораторная работа № 1. Однофакторные регрессионные модели и методы их построения | 4 |
| Лабораторная работа № 2. Модель множественной регрессии и методы ее построения | 22 |
| Лабораторная работа № 3. Гетероскедастичность модели и метод взвешенных наименьших квадратов | 38 |
| Лабораторная работа 4. Авторегрессионные процессы | 46 |
| Литература | 58 |
| Приложение 1 | 59 |
| Приложение 2 | 60 |
| Приложение 3 | 62 |
| Приложение 4 | 64 |

2. РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Задача 1. По территориям региона приводятся данные за 20XX г.

Таблица 1.1

| Номер региона | Среднедушевой прожиточный минимум в день одного трудоспособного, руб., x | Среднедневная заработная плата, руб., y |
|---------------|--|---|
| 1 | 85 | 142 |
| 2 | 89 | 148 |
| 3 | 87 | 142 |
| 4 | 79 | 154 |
| 5 | 89 | 164 |
| 6 | 113 | 195 |
| 7 | 67 | 139 |
| 8 | 98 | 167 |
| 9 | 82 | 152 |
| 10 | 87 | 162 |
| 11 | 86 | 155 |
| 12 | 117 | 173 |

Требуется:

1. Построить линейное уравнение парной регрессии y по x .
2. Рассчитать линейный коэффициент парной корреляции, коэффициент детерминации и среднюю ошибку аппроксимации.
3. Оценить статистическую значимость уравнения регрессии в целом и отдельных параметров регрессии и корреляции с помощью F -критерия Фишера и t -критерия Стьюдента.
4. Выполнить прогноз заработной платы y при прогнозном значении среднедушевого прожиточного минимума x , составляющем 107 % от среднего уровня.
5. Рассчитать и интерпретировать коэффициент эластичности.
6. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.
7. На одном графике отложить исходные данные и теоретическую прямую.

Все расчеты провести в *Excel* с использованием вышеприведенных формул и «Пакета анализа». Результаты, полученные по формулам и с помощью «Пакета анализа», сравнить между собой.

Решение:

1. Для определения параметров уравнения линейной регрессии строим расчетную таблицу 1.2.

Таблица 1.2

| № | x | y | y · x | x ² | y ² | \hat{y}_x | y – \hat{y}_x | (y – \hat{y}_x) ² | A _i |
|----------------|--------|--------|----------|----------------|----------------|-------------|-----------------|---------------------------------|----------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 85 | 142 | 12070 | 7225 | 20164 | 153,12 | –11,12 | 123,54 | 7,83 |
| 2 | 89 | 148 | 13172 | 7921 | 21904 | 156,91 | –8,91 | 79,41 | 6,02 |
| 3 | 87 | 142 | 12354 | 7569 | 20164 | 155,01 | –13,01 | 169,34 | 9,16 |
| 4 | 79 | 154 | 12166 | 6241 | 23716 | 147,42 | 6,58 | 43,28 | 4,27 |
| 5 | 89 | 164 | 14596 | 7921 | 26896 | 156,91 | 7,09 | 50,25 | 4,32 |
| 6 | 113 | 195 | 22035 | 12769 | 38025 | 179,69 | 15,31 | 234,49 | 7,85 |
| 7 | 67 | 139 | 9313 | 4489 | 19321 | 136,03 | 2,97 | 8,80 | 2,13 |
| 8 | 98 | 167 | 16366 | 9604 | 27889 | 165,45 | 1,55 | 2,40 | 0,93 |
| 9 | 82 | 152 | 12464 | 6724 | 23104 | 150,27 | 1,73 | 3,00 | 1,14 |
| 10 | 87 | 162 | 14094 | 7569 | 26244 | 155,01 | 6,99 | 48,82 | 4,31 |
| 11 | 86 | 155 | 13330 | 7396 | 24025 | 154,06 | 0,94 | 0,88 | 0,60 |
| 12 | 117 | 173 | 20241 | 13689 | 29929 | 183,48 | –10,48 | 109,89 | 6,06 |
| Итого | 1079 | 1893 | 172201 | 99117 | 301381 | 1893,37 | –0,37 | 874,10 | 54,64 |
| Средн. знач. | 89,92 | 157,75 | 14350,08 | 8259,75 | 25115,08 | 157,78 | – | 72,84 | 4,55 |
| σ | 13,20 | 15,17 | – | – | – | – | – | – | – |
| σ ² | 174,14 | 230,02 | – | – | – | – | – | – | – |

Находим параметры уравнения регрессии:

$$\hat{b}_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{14350,08 - 89,92 \cdot 157,75}{8259,75 - 89,92^2} = \frac{165,20}{174,14} = 0,949;$$

$$\hat{b}_0 = \bar{y} - \hat{b}_1 \cdot \bar{x} = 157,75 - 0,949 \cdot 89,92 = 72,45.$$

Получено уравнение регрессии $y = 72,45 + 0,949x$.

Параметр регрессии позволяет сделать вывод, что с увеличением среднедушевого прожиточного минимума на 1 руб. среднедневная заработная плата возрастает в среднем на 0,95 руб. (или 95 коп.).

После нахождения уравнения регрессии заполняем столбцы 7—10 табл. 1.2.

1. Тесноту линейной связи оценит коэффициент корреляции

$$r_{xy} = b_1 \frac{\sigma_x}{\sigma_y} = 0,949 \cdot \frac{13,20}{15,17} = 0,826.$$

Так как значение коэффициента корреляции больше 0,7, то это говорит о наличии весьма тесной линейной связи между признаками.

Коэффициент детерминации

$$D = r_{xy}^2 = 0,682.$$

Это означает, что 68,2 % вариации заработной платы (y) объясняется вариацией фактора x — среднедушевого прожиточного минимума.

Качество модели определяет средняя ошибка аппроксимации:

$$\bar{A} = \frac{1}{n} \sum A_i \times 100 \% = \frac{54,64}{12} = 4,55 \%.$$

Качество построенной модели оценивается как хорошее, поскольку \bar{A} не превышает 10 %.

Оценку статистической значимости уравнения регрессии в целом проведем с помощью F -критерия Фишера. Фактическое значение F -критерия составит

$$F_{расч} = \frac{r_{xy}^2}{1 - r_{xy}^2} (n - 2) = \frac{0,682}{1 - 0,682} \cdot 10 = 21,45.$$

Табличное значение критерия при пятипроцентном уровне значимости и степенях свободы $k_1 = 1$ и $k_2 = 12 - 2 = 10$ составляет $F_{табл} = 4,96$. Так как $F_{расч} = 21,45 > F_{табл} = 4,96$, то уравнение регрессии признается статистически значимым.

3. Оценку статистической значимости параметров регрессии и корреляции проведем с помощью t -статистики Стьюдента и путем расчета доверительного интервала каждого из параметров.

Табличное значение t -критерия для числа степеней свободы $df = n - m - 1 = 12 - 1 - 1 = 10$ и уровня значимости $\alpha = 0,05$ составит $t_{табл} = 2,23$.

Определим стандартные ошибки m_a , m_b , $m_{r_{xy}}$ (остаточная дисперсия

на одну степень свободы $S_{ост}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - m - 1} = \frac{874,10}{10} = 87,74$):

$$s_{b_0} = \sqrt{S_{ост}^2 \frac{\sum x_i^2}{n^2 \sigma_x^2}} = \sqrt{87,74 \cdot \frac{99117}{12^2 \cdot 174,14}} = 18,62;$$

$$s_{b_1} = \sqrt{\frac{S_{ocm}^2}{\sigma_x^2 n}} = \sqrt{\frac{87,74}{12 \cdot 174,14}} = 0,205;$$

$$s_{r_{xy}} = \sqrt{\frac{1 - r_{xy}^2}{n - 2}} = \sqrt{\frac{1 - 0,682}{12 - 2}} = 0,178.$$

Тогда

$$t_{b_0} = \frac{b_0}{s_{b_0}} = \frac{72,45}{18,62} = 3,89;$$

$$t_{b_1} = \frac{b_1}{s_{b_1}} = \frac{0,949}{0,205} = 4,63;$$

$$t_{r_{xy}} = \frac{r_{xy1}}{s_{r_{xy}}} = \frac{0,826}{0,178} = 4,64.$$

Фактические значения t -статистики превосходят табличное значение:

$$t_{b_0} = 3,89 > t_{табл} = 2,23;$$

$$t_{b_1} = 4,63 > t_{табл} = 2,23;$$

$$t_{r_{xy}} = 4,64 > t_{табл} = 2,23.$$

поэтому параметры \hat{b}_0 , \hat{b}_1 и r_{xy} не случайно отличаются от нуля, а статистически значимы.

Рассчитаем доверительные интервалы для параметров регрессии \hat{b}_0 и \hat{b}_1 . Для этого определим предельную ошибку для каждого показателя:

$$\Delta_{\hat{b}_0} = t_{табл} \cdot s_{b_0} = 2,23 \cdot 18,62 = 41,52;$$

$$\Delta_{\hat{b}_1} = t_{табл} \cdot s_{b_1} = 2,23 \cdot 0,205 = 0,457.$$

Доверительные интервалы:

$$\hat{b}_0 - \Delta_{\hat{b}_0} \leq b_0 \leq \hat{b}_0 + \Delta_{\hat{b}_0}, \quad \hat{b}_1 - \Delta_{\hat{b}_1} \leq b_1 \leq \hat{b}_1 + \Delta_{\hat{b}_1},$$

$$72,45 - 41,52 \leq b_0 \leq 72,45 + 41,52, \quad 0,949 - 0,457 \leq b_1 \leq 0,949 + 0,457,$$

$$30,92 \leq b_0 \leq 113,98; \quad 0,492 \leq b_1 \leq 1,406.$$

Анализ верхней и нижней границ доверительных интервалов приводит к выводу о том, что с вероятностью $p = 1 - \alpha = 0,95$ параметры \hat{b}_0 и \hat{b}_1 ,

находясь в указанных границах, не принимают нулевых значений, то есть являются статистически значимыми и существенно отличны от нуля.

4. Усредненное значение коэффициента эластичности

$$E = \hat{b}_1 \frac{\bar{x}}{\bar{y}} = 0,949 \cdot \frac{89,92}{157,75} = 0,54$$

показывает, на сколько процентов в среднем по совокупности изменится результат y от своей средней величины при изменении фактора x на 1 % от своего среднего значения.

5. Полученные оценки уравнения регрессии позволяют использовать его для прогноза. Если прогнозное значение прожиточного минимума составит $x_0 = \bar{x} \cdot 1,07 = 89,92 \cdot 1,07 = 96,21$ руб., тогда индивидуальное прогнозное значение заработной платы составит $\hat{y}_0 = 72,45 + 0,949 \cdot 96,21 = 163,76$ руб.

6. Ошибка прогноза составит

$$s_{\hat{y}_0} = \sqrt{S_{ост}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{n\sigma_x^2} \right)} = \sqrt{87,74 \left(1 + \frac{1}{12} + \frac{(96,21 - 89,92)^2}{12 \cdot 174,14} \right)} = 9,83.$$

Предельная ошибка прогноза, которая в 95 % случаев не будет превышена, составит

$$\Delta_{\hat{y}_0} = t_{табл} \cdot s_{\hat{y}_0} = 2,23 \cdot 9,83 = 21,93.$$

Доверительный интервал прогноза:

$$\begin{aligned} \hat{y}_0 - \Delta_{\hat{y}_0} &\leq \hat{y}_0 \leq \hat{y}_0 + \Delta_{\hat{y}_0}, \\ 163,76 - 21,93 &\leq \hat{y}_0 \leq 163,76 + 21,93, \\ 141,83 &\leq \hat{y}_0 \leq 185,69 \end{aligned}$$

Выполненный прогноз среднемесячной заработной платы является надежным ($p = 1 - \alpha = 1 - 0,05 = 0,95$) и находится в пределах от 141,83 руб. до 185,69 руб.