*ЛАБОРАТОРНАЯ РАБОТА №1*

По предприятиям легкой промышленности региона получена информация, характеризующая зависимость объема выпуска продукции (Y, млн. руб.) от объема капиталовложений (X, млн. руб.)

Требуется:

1. Построить диаграмму рассеяния для переменных «объёмы продаж» -Y и «индекс потребительских расходов» - X
2. Оценить тесноту взаимосвязи.
3. Оценить значимость вычисленного коэффициента парной корреляции.
4. Найти параметры уравнения линейной регрессии, дать экономическую интерпретацию коэффициента регрессии.
5. Вычислить остатки; найти остаточную сумму квадратов; оценить дисперсию остатков S2ε; построить график остатков.
6. Проверить выполнение предпосылок МНК.
7. Осуществить проверку значимости параметров уравнения регрессии с помощью t-критерия Стьюдента (α = 0,05)
8. Вычислить коэффициент детерминации, проверить значимость уравнения регрессии с помощью F- критерия Фишера (α = 0,05), найти среднюю относительную ошибку аппроксимации. Сделать вывод о качестве модели.
9. Осуществить прогнозирование среднего значения показателя Y при уровне значимости α = 0,1, если прогнозное значения фактора Х составит 80% от его максимального значения.
10. Представить графически: фактические и модельные значения Y, точки прогноза.
11. Составить уравнения нелинейной регрессии:
* гиперболической;
* степенной;
* показательной.

Привести графики построенных уравнений регрессии.

1. Для указанных моделей найти коэффициенты детерминации и средние относительные ошибки аппроксимации. Сравнить модели по этим характеристикам и сделать вывод.

Вариант 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 66 | 58 | 73 | 82 | 81 | 84 | 55 | 67 | 81 | 59 |
| Y | 133 | 107 | 145 | 162 | 163 | 170 | 104 | 132 | 159 | 116 |

Решение:

1.Диаграмма рассеяния имеет вид:

2. Средние значения случайных величин х и у:

$\overbar{х}=\frac{1}{n}\sum\_{i=1}^{n}x\_{i}=70.6$$\overbar{y}=\frac{1}{n}\sum\_{i=1}^{n}y\_{i}=139.1$

$$S\_{x}^{2}=\frac{1}{n-1}\sum\_{}^{}(x\_{i}-\overbar{x})^{2}=\frac{1102.4}{9}=122.49$$

$$S\_{y}^{2}=\frac{1}{n-1}\sum\_{}^{}(y\_{i}-\overbar{y})^{2}=\frac{5364.9}{9}=596.1$$

Рассчитаем стандартные ошибки случайных величин х и у:

$S\_{x}=\sqrt{\frac{1}{n-1}\sum\_{i=1}^{n}(x\_{i}-\overbar{x})^{2}}$=11.07

$S\_{y}=\sqrt{\frac{1}{n-1}\sum\_{i=1}^{n}(y\_{i}-\overbar{y})^{2}}$=24.42

Рассчитаем коэффициент корреляции:

$r\_{xy}$*=*=$\frac{\frac{1}{9}∙2418.4}{11.07∙24.42}=$ 0.994

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *y* | *x* | $$y\_{i}-\overbar{y}$$ | $$x\_{i}-\overbar{x}$$ | $$(y\_{i}-\overbar{y})(x\_{i}-\overbar{x})$$ | $$(x\_{i}-\overbar{x})^{2}$$ | $$(y\_{i}-\overbar{y})^{2}$$ |
| 1 | 133 | 66 | -6,1 | -4,6 | 28,06 | 21,16 | 37,21 |
| 2 | 107 | 58 | -32,1 | -12,6 | 404,46 | 158,76 | 1030,41 |
| 3 | 145 | 73 | 5,9 | 2,4 | 14,16 | 5,76 | 34,81 |
| 4 | 162 | 82 | 22,9 | 11,4 | 261,06 | 129,96 | 524,41 |
| 5 | 163 | 81 | 23,9 | 10,4 | 248,56 | 108,16 | 571,21 |
| 6 | 170 | 84 | 30,9 | 13,4 | 414,06 | 179,56 | 954,81 |
| 7 | 104 | 55 | -35,1 | -15,6 | 547,56 | 243,36 | 1232,01 |
| 8 | 132 | 67 | -7,1 | -3,6 | 25,56 | 12,96 | 50,41 |
| 9 | 159 | 81 | 19,9 | 10,4 | 206,96 | 108,16 | 396,01 |
| 10 | 116 | 59 | -23,1 | -11,6 | 267,96 | 134,56 | 533,61 |
| сумма | 1391 | 706 |  |  | 2418,4 | 1102,4 | 5364,9 |
| ср. Знач | 139,1 | 70,6 |  |  |  |  |  |

3. Оценим значимость коэффициента корреляции. Рассчитаем значение t–статистики:

t ==$\frac{0,994\sqrt{8}}{\sqrt{1-0,988}}$ =25,324

Полученное значение коэффициента корреляции значимо.

4. Построение линейной модели парной регрессии.

Определим линейный коэффициент парной корелляции:

$$r\_{xy}=\frac{\sum\_{}^{}(y\_{i}-\overbar{y})(x\_{i}-\overbar{x})}{\sqrt{\sum\_{}^{}(y\_{i}-\overbar{y})^{2}×\sum\_{}^{}(x\_{i}-\overbar{x})^{2}}}=\frac{2418.4}{\sqrt{5364.9∙1102.4}}=0.994$$

Уравнение линейной регрессии имеет вид: ŷ = a + b × x

Значения параметров a и b линейной модели определим, используя данные таблицы

b=

a=$\overbar{y}-b×\overbar{x}=$ 139.1-2.194\*70.6=-15.796

$\hat{y\_{1}}=$ -15.796+2.194\*66=129.008

$$\hat{y\_{2}}=-15.796+2.194\*58=111.456$$

$$\hat{y\_{3}}=-15.796+2.194\*73=144.366$$

$$\hat{y\_{4}}=-15.796+2.194\*82=164.112$$

$$\hat{y\_{5}}=-15.796+2.194\*81=161.918$$

$$\hat{y\_{6}}=-15.796+2.194\*84=168.5$$

$$\hat{y\_{7}}=-15.796+2.194\*55=104.874$$

$$\hat{y\_{8}}=-15.796+2.194\*67=131.202$$

$$\hat{y\_{9}}=-15.796+2.194\*81=161.918$$

$$\hat{y\_{10}}=-15.796+2.194\*59=113.65$$

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | y | x | y\*x | x\*x | $$\hat{y}$$ | ε | $$\left|ε/y\right|\*100\%$$ |
| 1 | 133 | 66 | 8778 | 4356 | 129,008 | 3,992 | 3,001504 |
| 2 | 107 | 58 | 6206 | 3364 | 111,456 | -4,456 | -4,16449 |
| 3 | 145 | 73 | 10585 | 5329 | 144,366 | 0,634 | 0,437241 |
| 4 | 162 | 82 | 13284 | 6724 | 164,112 | -2,112 | -1,3037 |
| 5 | 163 | 81 | 13203 | 6561 | 161,918 | 1,082 | 0,663804 |
| 6 | 170 | 84 | 14280 | 7056 | 168,5 | 1,5 | 0,882353 |
| 7 | 104 | 55 | 5720 | 3025 | 104,874 | -0,874 | -0,84038 |
| 8 | 132 | 67 | 8844 | 4489 | 131,202 | 0,798 | 0,604545 |
| 9 | 159 | 81 | 12879 | 6561 | 161,918 | -2,918 | -1,83522 |
| 10 | 116 | 59 | 6844 | 3481 | 113,65 | 2,35 | 2,025862 |
| сумма | 1391 | 706 | 100623 | 50946 |  | -0,004 | -0,52849 |
| ср. Знач | 139,1 | 70,6 | 10062,3 | 5094,6 |  |  | -0,05285 |

Уравнение линейной регрессии имеет вид:

$\hat{у}=$ -15,796+2,194\*х

Рассчитаем коэффициент детерминации:

$$R^{2}=r\_{yx}^{2}=0.988$$

Оценку значимости уравнения регрессии проведем с помощью критерия Фишера:

$$F=\frac{r\_{yx}^{2}}{1-r\_{yx}^{2}}×\left(n-2\right)=\frac{0.988}{1-0.988}×8=658.667$$

Уравнение регрессии статистически значимое.

Определим среднюю относительную ошибку:

$$\overbar{E\_{oтн}}=\frac{1}{n}×\sum\_{}^{}\frac{\left|E\_{i}\right|}{y}×100\%=\frac{-0.528}{10}=-0.0528\%$$

В среднем расчетные значения $\hat{у} $для линейной модели отличаются от фактических значений на -0,0528%.

5. Построение степенной модели.

Уравнение степенной модели имеет вид: $\hat{у}=а×х^{b}$

$$Произведем логарифмирование обеих частей уравнения:$$

lg $\hat{у} $= lg a + b lg x

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | y(t) | lg (y) | x(t) | lg (x) |
| 1 | 133.0 | 2.124 | 66 | 1.819 |
| 2 | 107.0 | 2.029 | 58 | 1.763 |
| 3 | 145.0 | 2.161 | 73 | 1.863 |
| 4 | 162.0 | 2.209 | 82 | 1.914 |
| 5 | 163.0 | 2.212 | 81 | 1.908 |
| 6 | 170.0 | 2.231 | 84 | 1.924 |
| 7 | 104.0 | 2.017 | 55 | 1.741 |
| 8 | 132.0 | 2.121 | 67 | 1.826 |
| 9 | 159.0 | 2.201 | 81 | 1.908 |
| 10 | 116.0 | 2.064 | 59 | 1.771 |
| итого | 1391 | 21.369 | 706 | 18.437 |
| ср. знач. | 139.1 | 2.137 | 70.6 | 1.844 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | y | Y | x | X | YX | X\*X | $$\hat{y}$$ | ε | $$\left|ε/y\right|\*100\%$$ | $$ε^{2}$$ |
| 1 | 133 | 2,123852 | 66 | 1,819544 | 3,864441 | 3,31074 | 127,732 | 5,268 | 3,960902 | 27,75182 |
| 2 | 107 | 2,029384 | 58 | 1,763428 | 3,578672 | 3,109678 | 108,681 | -1,681 | -1,57103 | 2,825761 |
| 3 | 145 | 2,161368 | 73 | 1,863323 | 4,027326 | 3,471972 | 144,885 | 0,115 | 0,07931 | 0,013225 |
| 4 | 162 | 2,209515 | 82 | 1,913814 | 4,2286 | 3,662683 | 167,547 | -5,547 | -3,42407 | 30,76921 |
| 5 | 163 | 2,212188 | 81 | 1,908485 | 4,221927 | 3,642315 | 164,997 | -1,997 | -1,22515 | 3,988009 |
| 6 | 170 | 2,230449 | 84 | 1,924279 | 4,292007 | 3,702851 | 172,671 | -2,671 | -1,57118 | 7,134241 |
| 7 | 104 | 2,017033 | 55 | 1,740363 | 3,51037 | 3,028862 | 101,701 | 2,299 | 2,210577 | 5,285401 |
| 8 | 132 | 2,120574 | 67 | 1,826075 | 3,872327 | 3,334549 | 130,156 | 1,844 | 1,39697 | 3,400336 |
| 9 | 159 | 2,201397 | 81 | 1,908485 | 4,201333 | 3,642315 | 164,997 | -5,997 | -3,7717 | 35,96401 |
| 10 | 116 | 2,064458 | 59 | 1,770852 | 3,65585 | 3,135917 | 111,029 | 4,971 | 4,285345 | 24,71084 |
| Сумма | 1391 | 21,37022 | 706 | 18,43865 | 39,45285 | 34,04188 |  | -3,396 |  | 141,8429 |

Y=lg $\hat{у}$, X=lg x, A=lg a.

Уравнение примет вид: Y=A+bX – линейное уравнение регрессии.

$$b=\frac{\overbar{Y∙X}- \overbar{Y} ∙\overbar{X}}{\overbar{X^{2}}- \overbar{X}^{2}}=\frac{3.9449-2.1369\*1.8437}{3.4036-1.8437\*1.8437}=\frac{0.005}{0.004}=1.25 $$

$$A=\overbar{Y}-b\overbar{X}=2.1369-1.25\*1.8437=-0.168$$

Уравнение регрессии: Y=-0,168+1,25Х

$$\hat{у}=10^{-0,168}∙х^{1,25}$$

$$\hat{у}=0,679∙х^{1,25}$$

Индекс корреляции:

$$ρ\_{yx}=\sqrt{1-\frac{\sum\_{}^{}(y-\hat{y})^{2}}{\sum\_{}^{}(y-\overbar{y})^{2}}}=0.994$$

Связь между показателем у и фактором х можно считать достаточно сильной.

Коэффициент детерминации: 0,989

$$R^{2}=ρ\_{yx}^{2}=0.994^{2}=0.989$$

F критерий Фишера:

$$F=\frac{R^{2}}{1-R^{2}}×\left(n-2\right)=\frac{0.989}{1-0.989}×8=719.273$$

Cредняя относительная ошибка:

$$\overbar{E\_{oтн}}=\frac{0,37}{10}=0,037\%$$

В среднем расчетные значения$ \hat{у}$ для степенной модели отличаются от фактических на 0,037%.

6. Построение показательной функции

Уравнение показательной кривой: $\hat{у}=abx$

lg $\hat{у}=$ lg a +x lg b

Y= lg $\hat{у}$, B=lg b, A=lg a.

Получим линейное уравнение регрессии: Y=A+Bx

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | y | Y | x | Yx | x\*x | У-$\overbar{У}$ |
| 1 | 133 | 2,124 | 66 | 140,184 | 4356 | -0,0129 |
| 2 | 107 | 2,029 | 58 | 117,682 | 3364 | -0,1079 |
| 3 | 145 | 2,161 | 73 | 157,753 | 5329 | 0,0241 |
| 4 | 162 | 2,209 | 82 | 181,138 | 6724 | 0,0721 |
| 5 | 163 | 2,212 | 81 | 179,172 | 6561 | 0,0751 |
| 6 | 170 | 2,231 | 84 | 187,404 | 7056 | 0,0941 |
| 7 | 104 | 2,017 | 55 | 110,935 | 3025 | -0,1199 |
| 8 | 132 | 2,121 | 67 | 142,107 | 4489 | -0,0159 |
| 9 | 159 | 2,201 | 81 | 178,281 | 6561 | 0,0641 |
| 10 | 116 | 2,064 | 59 | 121,776 | 3481 | -0,0729 |
| итого | 1391 | 21,369 | 706 | 1516,432 | 50946 |  |
| Ср. знач | 139,1 | 2,1369 | 70,6 |  |  |  |
| $$\left(Y-\overbar{Y}\right)^{2}$$ | x-$\overbar{x}$ | $$\left(x-\overbar{x}\right)^{2}$$ | $$\hat{y}$$ | $$(y-\hat{y})^{2}$$ | ε | $$\left|ε/y\right|\*100\%$$ |
| 0,000166 | -4,6 | 21,16 | 125,298 | 7,702 | 59,3208 | 5,791 |
| 0,011642 | -12,6 | 158,76 | 110,356 | -3,356 | 11,26274 | -3,136 |
| 0,000581 | 2,4 | 5,76 | 140,023 | 4,977 | 24,77053 | 3,4324 |
| 0,005198 | 11,4 | 129,96 | 161,526 | 0,474 | 0,224676 | 0,2926 |
| 0,00564 | 10,4 | 108,16 | 158,982 | 4,018 | 16,14432 | 2,465 |
| 0,008855 | 13,4 | 179,56 | 166,736 | 3,264 | 10,6537 | 1,92 |
| 0,014376 | -15,6 | 243,36 | 105,224 | -1,224 | 1,498176 | -1,177 |
| 0,000253 | -3,6 | 12,96 | 127,303 | 4,697 | 22,06181 | 3,5583 |
| 0,004109 | 10,4 | 108,16 | 158,982 | 0,018 | 0,000324 | 0,0113 |
| 0,005314 | -11,6 | 134,56 | 112,121 | 3,879 | 15,04664 | 3,344 |
| 0,056135 |  | 1102,4 |  |  | 160,9837 | 16,5016 |
|  |  |  |  |  |  | 1,6502 |

$$B=\frac{\overbar{Y∙x}- \overbar{Y }∙\overbar{x}}{\overbar{x^{2}} ∙ \overbar{x}^{2}}=\frac{151.643-2.137\*70.6}{5094.6-70.6\*70.6}=0.007$$

$$A=\overbar{Y}-B∙\overbar{x}=2.137-0.007\*70.6=1.643$$

Уравнение будет иметь вид: Y=1.643+0.007x

$$\hat{y}=10^{1.643}\*\left(10^{0.007}\right)^{x}=43.95\*1.016^{x}$$

Индекс корреляции: $ρ\_{yx}=\sqrt{1-\frac{\sum\_{}^{}(y-\hat{y})^{2}}{\sum\_{}^{}(y-\overbar{y})^{2}}}$ =$\sqrt{1-\frac{16.098}{5364.9}}=0.998$

Индекс детерминации: $R^{2}=P\_{yx}^{2}$=0.996

F критерий Фишера:

$$F=\frac{R^{2}}{1-R^{2}}×\left(n-2\right)=\frac{0.996}{1-0.996}\*8=1992$$

Cредняя относительная ошибка:

$$\overbar{E\_{oтн}}=\frac{16.5016}{10}=1.6502\%$$

В среднем расчетные значения$ \hat{у}$ для показательной функции отличаются от фактических на 1,6502%.

7. Построение гиперболической функции

Уравнение гиперболической функции: $\hat{у}=а$+b/x

X=1/x

$\hat{у}=а$+bX

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | y | x | X | yX | X\*X | У-$\overbar{У}$ |
| 1 | 133 | 66 | 0,015152 | 2,015152 | 0,00023 | -6,1 |
| 2 | 107 | 58 | 0,017241 | 1,844828 | 0,000297 | -32,1 |
| 3 | 145 | 73 | 0,013699 | 1,986301 | 0,000188 | 5,9 |
| 4 | 162 | 82 | 0,012195 | 1,97561 | 0,000149 | 22,9 |
| 5 | 163 | 81 | 0,012346 | 2,012346 | 0,000152 | 23,9 |
| 6 | 170 | 84 | 0,011905 | 2,02381 | 0,000142 | 30,9 |
| 7 | 104 | 55 | 0,018182 | 1,890909 | 0,000331 | -35,1 |
| 8 | 132 | 67 | 0,014925 | 1,970149 | 0,000223 | -7,1 |
| 9 | 159 | 81 | 0,012346 | 1,962963 | 0,000152 | 19,9 |
| 10 | 116 | 59 | 0,016949 | 1,966102 | 0,000287 | -23,1 |
| итого | 1391 | 706 | 0,144939 | 19,64817 | 0,00215 |  |
| Ср. знач | 139,1 | 70,6 |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| $$\left(Y-\overbar{Y}\right)^{2}$$ | $$\hat{y}$$ | ε | $$(y-\hat{y})^{2}$$ | $$\left|ε/y\right|\*100\%$$ |
| 37,21 | 132,021 | 0,979 | 0,958441 | 0,73609 |
| 1030,41 | 109,316 | -2,316 | 5,363856 | -2,16449 |
| 34,81 | 147,806 | -2,806 | 7,873636 | -1,93517 |
| 524,41 | 164,141 | -2,141 | 4,583881 | -1,3216 |
| 571,21 | 162,505 | 0,495 | 0,245025 | 0,303681 |
| 954,81 | 167,296 | 2,704 | 7,311616 | 1,590588 |
| 1232,01 | 99,098 | 4,902 | 24,0296 | 4,713462 |
| 50,41 | 134,478 | -2,478 | 6,140484 | -1,87727 |
| 396,01 | 162,505 | -3,505 | 12,28503 | -2,2044 |
| 533,61 | 112,491 | 3,509 | 12,31308 | 3,025 |
| 5364,9 | 1391,657 | -0,657 | 81,10465 | 0,865882 |

$$b=\frac{\overbar{y∙X}- \overbar{y} ∙\overbar{X}}{\overbar{X^{2}}- \overbar{X}^{2}}=\frac{1.9648-139.1\*0.0145}{0.000215-0.0145\*0.0145}=-10864.583 $$

$$a=\overbar{y}-b\overbar{X}=139.1+10864.583\*0.0145=296.636$$

Уравнение гиперболической модели: $\hat{у}=296.636-{10864.583}/{x}$

$ρ\_{yx}=\sqrt{1-\frac{\sum\_{}^{}(y-\hat{y})^{2}}{\sum\_{}^{}(y-\overbar{y})^{2}}}$ =$\sqrt{1-\frac{81.105}{5364.9}}=0.992$

$$R^{2}=ρ\_{yx}^{2}=0.992^{2}=0.985$$

$$F=\frac{R^{2}}{1-R^{2}}×\left(n-2\right)=\frac{0.985}{1-0.985}\*8=525.333$$

Cредняя относительная ошибка:

$$\overbar{E\_{oтн}}=\frac{0.866}{10}=0.0866\%$$

В среднем расчетные значения$ \hat{у}$ для гиперболической модели отличаются от фактических значений на 0.0866%.

Для выбора лучшей модели построим сводную таблицу результатов:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Коэффициент детерминации $R^{2}$ | F критерий Фишера | Индекс корреляции $ρ\_{yx}$ | Cредняя относительная ошибка $\overbar{E\_{oтн}}$ |
| Линейная модель | 0.988 | 658.667 | 0.994 | -0.0528 |
| Степенная модель | 0.989 | 719.273 | 0.994 | 0.037 |
| Показательная модель | 0.996 | 1992 | 0.998 | 1.6502 |
| Гиперболическая модель | 0.985 | 525.333 | 0.992 | 0.0866 |

Все модели имеют примерно одинаковые характеристикиб но большее значение критерия Фишера и большее значение коэффициента детерминации имеет гиперболическая модель. Ее можно взять в качестве лучшей для построения прогноза.