**Задание**

Экономист, изучая зависимость уровня Y (тыс. руб.) издержек oбpaщения от объема X (тыс. руб.) товарооборота, обследовал по 10 магазинов, торгующих одинаковым ассортиментом товаров в 5 районах. Полученные данные отражены в таблице 1.

• найти коэффициенты корреляции между X и Y;

• построить регрессионные функции линейной зависимости Y = a + b \* X фактора Y от фактора X и исследовать их на надежность по критерию Фишера при уровне значимости 0,05;

• определить надежность коэффициентов регрессии по критерию Стьюдента:

• используя полученное уравнение линейной регрессии, оценить ожидаемое среднее значение признака Y при X = 130 тыс. руб.

|  |  |
| --- | --- |
| Объем товарооборота,тыс. руб. | Уровень издержек обращения,тыс. руб. |
| 60 | 2,9 |
| 90 | 7,1 |
| 1 6 0 | 1 2 |
| 80 | 6,3 |
| 10 5 | 7 |
| 120 | 8,4 |
| 70 | 4,8 |
| 130 | 11,2 |
| 1 1 0 | 7. 6 |
| 140 | 10,6 |

**Решение:**

Обозначим: X – объем товарооборота (тыс. руб.); Y – издержки обращения.

 Для проведения расчетов построим вспомогательную таблицу 1.1.

Таблица 1.1

 Предположим, что связь между объемом товарооборота и издержками обращения линейная. Для подтверждения нашего предположения построим поле корреляции (рис.1.1).

Рис. 1.1. – Поле корреляции

Уравнение прямой линии y = a + bx

Решим систему уравнений:

$$\left\{\begin{array}{c}10a+b\sum\_{}^{}x= \sum\_{}^{}y\\a\sum\_{}^{}x+b\sum\_{}^{}x^{2}= \sum\_{}^{}yx\end{array}\right.$$

 Подставим рассчитанные в таблице значения:

$$\left\{\begin{array}{c}10a+1065b=78\\1065a +122625b= 9092\end{array}\right.$$

 Выразим из первого уравнения a = $\frac{78-1065b}{10}$ и подставим во второе:

1065 $\frac{78-1065b}{10}$ $+122625b= 9092$

83070 – 1134225b + 1226250b = 90920

92025b = 7850

b = 0,085 ≈ 0,09

a = $\frac{78-1065b}{10}$ = a = $\frac{78-1065\*0,085}{10}$ = - 1,25

Получаем уравнение y = -1,25 + 0,09x

 Вывод:

 При увеличении товарооборота на 1 тыс. руб. издержки обращения увеличатся на 90 рублей.

Дисперсии будут равны:

$σ\_{x}^{2}$ = $(x^{2})\_{ср}$ - $x\_{ср}^{2}$ = 12263 – 11342,25 = 920,75;

$σ\_{x}$ = 30,34;

$σ\_{y}^{2}$ = $(y^{2})\_{ср}$ - $y\_{ср}^{2}$ = 68,07 – 60,68 = 7,39;

$σ\_{y}$ = 2,72.

Тесноту линейной связи оценивает коэффициент корреляции:

$r\_{xy}$ = b $\frac{σ\_{x}}{σ\_{y}}$ = 0,085 $\frac{30,34}{2,72}$ = 0,948;

R = $r\_{xy}^{2}$ = 0,8987.

 Вывод:

 Это означает, что 89,87% вариации издержек обращения объясняется вариацией фактора – объема товарооборота.

Качество модели определяет средняя ошибка аппроксимации:

$A\_{ср}$ = $\frac{1}{n}$ = $\sum\_{}^{}A\_{i}$ = 9,53%.

 Вывод:

 Качество построенной модели оценивается как хорошее, так как средняя ошибка аппроксимации не превышает 10%.

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

$F\_{факт}$ = $\frac{r\_{xy}^{2}}{1- r\_{xy}^{2}}$ (n – 2) = $\frac{0.8987}{1-0.8987}$ \* 8 = 70,97.

 Табличное значение критерия при пятипроцентном уровне значимости и степенях свободы $k\_{1}$ = 1 и $k\_{2}$ = 10 – 2 = 8 составляет $F\_{табл}$ = 5,32.

 Вывод:

 Так как $F\_{факт}$ = 70,97 > $F\_{табл}$ = 5,32 , то уравнение регрессии признается статистически значимым.

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

 Табличное значение -критерия для числа степеней свободы

df = n – 2 = 10 – 2 = 8 и  составит $t\_{табл}$ = 2,306.

Остаточная дисперсия:  = $\frac{5,06}{10-1-1}$ = 0,63; $S\_{ост}$ = 0,79

Определим случайные ошибки , , :

$m\_{a}$ = $\frac{S\_{ост}\sqrt{\sum\_{}^{}x^{2}}}{nσ\_{x}}$ = $\frac{0,79\sqrt{122625}}{10\*30,34}$ = 0,91;

$m\_{b}$ = $\frac{S\_{ост}}{σ\_{x}\sqrt{n}}$ = $\frac{0,79}{30,34\sqrt{10}}$ = 0,008;

$m\_{r\_{xy}}$ = $\sqrt{\frac{1-r\_{xy}^{2}}{n-2}}$ = $\sqrt{\frac{0.1013}{8}}$ = 0,113.

Тогда

$t\_{a}$ = $\frac{a}{m\_{a}}$ = $\frac{1,25}{0,91}$ = 1,37;

$t\_{b}$ = $\frac{b}{m\_{b}}$ = $\frac{0,09}{0,008}$ = 11,25;

$t\_{r\_{xy}}$ = $\frac{r\_{xy}}{m\_{r\_{xy}}}$ = $\frac{0,948}{0,113}$ = 8,39.

Вывод:

Фактическое значение -статистики для параметра a меньше табличного значения $t\_{a}$ = 1,37 < $t\_{табл}$ = 2,306, поэтому параметр a статистически незначим.

Фактические значения -статистики параметров  и  превосходят табличное значение $t\_{b}$ = 11,25 > $t\_{табл}$ = 2,306; $t\_{r\_{xy}}$ = 8,39 > $t\_{табл}$ = 2,306, поэтому параметры  и  статистически значимы.

Полученные оценки уравнения регрессии позволяют использовать его для прогноза. Если прогнозное значение объема товарооборота составит 130 тыс. руб., тогда прогнозное значение издержек обращения составит: $y\_{p}$ = -1,25 + 0,09 \* 130 = 10,45 тыс. руб.