# ***Раздел. Элементы теории корреляции***

# ***Образец решения***

Экономист, изучая зависимость производительности труда *Y* (т/ч) от уровня механизации работ *X* (%), обследовал 10 однотипных предприятий и получил следующие данные (табл.).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 53 | 31 | 77 | 60 | 37 | 69 | 47 | 54 | 66 | 40 |
|  | 37 | 25 | 49 | 40 | 29 | 45 | 33 | 38 | 42 | 30 |

Полагая, что между признаками *X* и *Y* имеет место линейная корреляционная связь, определите выборочное уравнение линейной регрессии и выборочный коэффициент линейной корреляции. Постройте диаграмму рассеяния и линию регрессии. Сделайте вывод о направлении и тесноте связи между *X* и *Y*.

**Решение**

Построим диаграмму рассеяния. Для этого на плоскости *xOy* отметим точки с координатами (*xi ; yi*).

По диаграмме рассеяния видно, что точки (*xi ; yi*) группируются около некоторой прямой. Поэтому выборочное уравнение линейной регрессии будем искать в виде *y = a∙x+b*. Параметры *a* и *b* найдем методом наименьших квадратов. Составим систему нормальных уравнений:



Вспомогательные вычисления проведем в таблице 8:

Таблица 8



Итак, система нормальных уравнений имеет вид:

.

Решим её методом Крамера. Определитель системы

.

.

.

, .

Выборочное уравнение линейной регрессии имеет вид *y=*0,506819*∙x+*9,73586.

Чтобы построить линию регрессии найдем координаты двух точек, принадлежащих прямой *y=*0,506819*∙x+*9,73586.

При *x*=35 *y*=0,506819∙35+9,73586=27,474529≈27,5.

При *x*=75 *y*=0,506819∙75+9,73586=47,747292≈47,7.

Линия регрессии – прямая, проходящая через точки (35; 27,5) и (75;47,7).

Выборочный коэффициент линейной корреляции найдем по формуле

, где  – наблюдавшиеся значения признаков *X* и *Y*;  – объём выборки;  – выборочные средние;  – выборочные среднеквадратические отклонения.

. .

.

.

.

Так как выборочный коэффициент линейной корреляции , то корреляция положительная, т. е. с возрастанием *x* возрастает и *y*. Так как  очень близко к единице, то связь между признаками *x* и *y* тесная.

**Ответ:** уравнение регрессии *y=*0,506819*∙x+*9,73586; связь между признаками *x* и *y* тесная, положительная.

# ***Вариант 1.***

Экономист, изучая зависимость выработки У (тыс. руб.) на одного работника от величины товарооборота магазина Х (тыс. руб.) за отчетный период обследовал семь магазинов и получил следующие данные.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Х* | 110 | 140 | 100 | 140 | 85 | 150 | 95 | 170 | 80 | 120 |
| *Y* | 9.5 | 13.0 | 10.0 | 12.0 | 8.0 | 13.0 | 8.0 | 14.0 | 6.5 | 11.0 |

Полагая, что между признаками Х и У имеет место линейная корреляционная связь, определить выборочное уравнение линейной регрессии и выборочный коэффициент линейной корреляции. Построить диаграмму рассеяния и линию регрессии. Сделать вывод о направлении и тесноте связи между Х и У. Используя полученное уравнение линейной регрессии, оценить х0=200 тыс. руб.

***Решение*.**

По данным наблюдений выборочное уравнение прямой линии среднеквадратической регрессии ****** *на*** представляется в виде:



Неизвестные параметры и находим из системы уравнений:



Для вычисления параметров  и  составим расчетную таблицу:



Тогда предыдущая система примет вид:



Решая эту систему получаем: и 

Уравнение регрессии  на  имеет вид:.

Построим диаграмму рассеяния и линию регрессии:

Точки А(90; 7,2) и В(150;12,85) находим из уравнения регрессии. Далее вычисляем:

, .

.

.

Находим коэффициент корреляции по формуле:



Итак, *rb*=0,97>0 => корреляция положительная, т.е. с возрастанием ** возрастает и **. Теснота связи достаточно большая, так как 

Зная уравнение регрессии, можно вычислить предполагаемую выработку от величины товарооборота при *X*0 = 200 тыс. руб.

:*y =*0,08*∙*200+0,8516,85

*Вывод*: зависимость выработки от величины товарооборота *X*0 = 200 тыс. руб. составит примерно 17 тысяч рублей.

***Вариант 2.***

Экономист, изучая зависимость выработки У (тыс. руб.) на одного работника от величины товарооборота магазина Х (тыс. руб.) за отчетный период обследовал семь магазинов и получил следующие данные.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Х* | 140 | 100 | 170 | 130 | 180 | 130 | 100 | 150 | 110 | 200 |
| *Y* | 8.0 | 9.0 | 11.0 | 8.0 | 11.0 | 7.5 | 5.0 | 10.0 | 7.5 | 11.0 |

Полагая, что между признаками Х и У имеет место линейная корреляционная связь, определить выборочное уравнение линейной регрессии и выборочный коэффициент линейной корреляции. Построить диаграмму рассеяния и линию регрессии. Сделать вывод о направлении и тесноте связи между Х и У. Используя полученное уравнение линейной регрессии, оценить Х0=210 тыс. руб.

***Решение*.**

Построим диаграмму рассеяния. Для этого на плоскости *xOy* отметим точки с координатами (*xi ; yi*).

По диаграмме рассеяния видно, что точки (*xi ; yi*) группируются около некоторой прямой. Поэтому выборочное уравнение линейной регрессии будем искать в виде



Параметры  и  найдем методом наименьших квадратов. Составим систему нормальных уравнений:



Для вычисления параметров  и  составим расчетную таблицу:



Итак, система нормальных уравнений имеет вид:



Решая эту систему, получаем: и .

Выборочное уравнение линейной регрессии имеет вид:.

Чтобы построить линию регрессии найдем координаты двух точек, принадлежащих прямой .

Линия регрессии – прямая, проходящая через точки А(100; 6,838) и В(200;11,623).

Выборочный коэффициент линейной корреляции найдем по формуле

, где  – наблюдавшиеся значения признаков *X* и *Y*;  – объём выборки;  – выборочные средние;  – выборочные среднеквадратические отклонения.

, .

.

.

Находим коэффициент корреляции по формуле:



Итак, *rb*=0,83>0 => корреляция положительная, т.е. с возрастанием ** возрастает и **. Теснота связи достаточно большая, так как 

Зная уравнение регрессии, можно вычислить предполагаемую выработку от величины товарооборота при *X*0 = 210 тыс. руб.

*y =*0,047855*∙*210+2,052431612,1.

*Вывод*: зависимость выработки от величины товарооборота *X*0 = 210 тыс. руб. составит примерно 12,1 тысяч рублей.

**Раздел. Оценки параметров распределения**

**Образец решения**

Из новорожденных детей, родившихся в Тираспольском роддоме в течение месяца, случайным образом отобраны 20 детей, рост которых составлял соответственно 50, 51, 48, 52, 50, 49 50, 47, 50, 51, 49, 50, 52, 48, 49, 51, 52, 48, 47, 51 сантиметр**.** Требуется: 1) определить выборочную среднюю , выборочную Dв и исправленную S2 дисперсии; 2) полагая, что распределение признака Х описывается нормальным законом распределения, найдите доверительный интервал для среднего роста а у обследуемых детей на уровне заданной надежности  = 0,999.

**Решение.**

1) Определим выборочную среднюю , выборочную Dв и исправленную S2 дисперсии. Вспомогательные расчеты проведем в таблице.

**

Выборочная средняя .

Выборочная дисперсия Dв.

Исправленная дисперсия .

2) Полагая, что распределение признака Х описывается нормальным законом распределения, найдём доверительный интервал для среднего роста а у обследуемых детей на уровне заданной надежности 

 = 0,999. Доверительный интервал найдём по формуле:

, где . по таблице по заданным n=20  = 0,999 находим t= 3,85.

Итак,  и доверительный интервал



Ответ: Выборочная средняя ; выборочная дисперсия Dв; исправленная дисперсия s2.

С надёжностью 0,999 средний рост а у обследуемых детей

заключен в доверительном интервале 48,37 < а < 51,13.

**Вариант 1.**

При изучении физико-механических свойств обувных кож было испытано n образцов и получены следующие значения предела прочности на разрыв *Х н/мм*.

Требуется:

а) определить выборочную среднюю, выборочную дисперсию, исправленную дисперсию;

б) полагая, что распределение величины *х* описывается нормальным законом, найти доверительный интервал для среднего предела прочности *a* этой кожи на уровне надежности γ= 0,95.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ образца** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|  | 14,6 | 19,5 | 20,0 | 16,8 | 19,4 | 17,1 | 18,2 | 17,5 |

**Решение.**

Статистическое распределение имеет вид:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 14.6 | 16.8 | 17.1 | 17.5 | 18.2 | 19.4 | 19.5 | 20.0 |
| *ni* | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |

Выборочная средняя .

Выборочная дисперсия Dв.

Исправленная дисперсия .

2) Полагая, что распределение признака *Х* описывается нормальным законом распределения, найдём доверительный интервал для среднего предела прочности *a* этой кожи на уровне надежности γ= 0,95.

Доверительный интервал найдём по формуле:

, где . по таблице по заданным n=8  = 0,95 находим t= 2,306.

Итак,  и доверительный интервал



Ответ: Выборочная средняя ; выборочная дисперсия Dв; исправленная дисперсия s2.

С надёжностью 0,95 средний предел прочности *a* этой кожи

заключен в доверительном интервале 13,3 < а < 18,1.

**Вариант 2.**

При изучении физико-механических свойств обувных кож было испытано n образцов и получены следующие значения предела прочности на разрыв *Х н/мм* .

Требуется:

а) определить выборочную среднюю, выборочную дисперсию, исправленную дисперсию;

б) полагая, что распределение величины *Х* описывается нормальным законом, найти доверительный интервал для среднего предела прочности *a* этой кожи на уровне надежности γ.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ образца** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|  | 23,3 | 19,4 | 20,1 | 24,3 | 22,8 | 18,0 | 17,5 | 17,1 |

**Решение.**

Статистическое распределение имеет вид:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 20.1 | 21 | 22,4 | 24,8 | 25,1 | 25,5 | 26,3 | 27,1 |
| *ni* | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Выборочная средняя .

Выборочная дисперсия Dв.

Исправленная дисперсия .

2) Полагая, что распределение признака *х* описывается нормальным законом распределения, найдём доверительный интервал для среднего предела прочности *a* этой кожи на уровне надежности  = 0,95. Доверительный интервал найдём по формуле:

, где . по таблице по заданным n=8  = 0,95 находим t= 2,306.

Итак,  и доверительный интервал



Ответ: Выборочная средняя ; выборочная дисперсия Dв; исправленная дисперсия s2.

С надёжностью 0,95 среднего предела прочности *a* этой кожи заключен

в доверительном интервале 21,95< а < 26,12.