**Методичні рекомендації**

Для проведення регресійного аналізу та прогнозування необхідно:

1) **побудувати графік** вихідних даних і спробувати візуально, наближено визначити характер залежності;

2) **обрати вид функції** регресії, яка може описувати зв'язок вихідних даних;

3) **визначити чисельні коефіцієнти** функції регресії методом найменших квадратів;

4) **оцінити силу** знайденої регресійної залежності на основі коефіцієнта детермінації ;

5) **зробити прогноз** (при ) або зробити висновок про неможливість прогнозування з допомогою знайденої регресійної залежності. При цьому не рекомендується використовувати модель регресії для тих значень незалежного параметра X, які не належать інтервалу, визначеного у вихідних даних.

**Довідкова информацця по технології роботи з режимом "Регресія" надбудови Пакет аналізу MS Excel**

Режим роботи "Регресія" служить для розрахунку параметрів рівняння лінійної регресії та перевірки її адекватності досліджуваному процесу.

Для вирішення задачі регресійного аналізу в MS Excel вибираємо у меню **Сервіс** команду **Аналіз даних** та інструмент аналізу "**Регресія**".

У діалоговому вікні задаємо наступні параметри:

1. *Вхідний інтервал Y* - це діапазон даних по результативному ознакою. Він повинен складатися з одного стовпця.

2. *Вхідний інтервал X* - це діапазон клітинок, що містять значення факторів (незалежних змінних). Число вхідних діапазонів (стовпців) має бути не більше 16.

3. *Прапорець Мітки*, встановлюється в тому випадку, якщо в першому рядку діапазону варто заголовок.

4. *Прапорець Рівень надійності* активізується, якщо в поле, що знаходиться поруч з ним необхідно ввести рівень надійності, відмінний від встановленого за замовчуванням. Використовується для перевірки значущості коефіцієнта детермінації  і коефіцієнтів регресії.

5. *Константа нуль*. Даний прапорець необхідно встановити, якщо лінія регресії має пройти через початок координат ().

6. *Вихідний інтервал/ Новий робочий аркуш/ Нова робоча книга* – вказати адресу верхньої лівої комірки вихідного діапазону.

7. Прапорці в розділі *Залишки* встановлюються, якщо необхідно включити у вихідний діапазон відповідні стовпці або графіки.

8. Прапорець Графік нормальної ймовірності необхідно зробити активним, якщо потрібно вивести на лист точковий графік залежності спостережуваних значень Y від автоматичних інтервалів персентилей.

Після натискання кнопки ОК у вихідному діапазоні отримуємо звіт.

**Лабораторна робота 1**

Завдання: Деяка фірма займається поставками різних вантажів на короткі відстані всередині міста. Оцінити вартість таких послуг, залежну від витрачається на поставку часу. В якості найважливішого чинника, що впливає на час поставки, вибрано пройдену відстань. Були зібрані вихідні дані про десять постачаннь (таблиця 1)

Таблиця 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Відстань, миль | 3,5 | 2,4 | 4,9 | 4,2 | 3,0 | 1,3 | 1,0 | 3,0 | 1,5 | 4,1 |
| Час, хвилини | 16 | 13 | 19 | 18 | 12 | 11 | 8 | 14 | 9 | 16 |

Визначте характер залежності між відстанню і витраченим часом, використовуючи майстер діаграм MS Еxcel, проаналізуйте застосовність методу найменших квадратів, побудуйте рівняння регресії, використовуючи МНК, проаналізуйте силу регресійної зв'язку. Проведіть регресійний аналіз, використовуючи режим роботи "Регресія" в MS Еxcel і порівняйте з результатами, отриманими раніше. Зробіть прогноз часу поїздки на 2 милі. Розрахувати та побудувати графічно міру помилки регресійної моделі використовуючи табличний процесор Excel.

Рішення

На графіку будуємо вихідні дані по десяти поїздок.



Рис.1. Графік вихідних даних і передбачувана лінія регресії

Крім відстані на час поставки впливають пробки на дорогах, час доби, дорожні роботи, погода, кваліфікація водія, вид транспорту. Побудовані точки не знаходяться точно на лінії, що обумовлено зазначеними вище чинниками. Але ці точки зібрані навколо прямої лінії, тому можна припустити лінійну зв'язок між параметрами. Всі вихідні точки рівномірно розподілені вздовж передбачуваної прямої лінії, що дозволяє застосувати метод найменших квадратів.

Обчислимо суми, необхідні для розрахунку коефіцієнтів рівняння лінійної регресії та коефіцієнта детермінації  за допомогою допоміжної таблиці (таблиця 2).

Таблица 2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 3,5 | 16 | 12,25 | 56,00 | 15,223 | 2,634129 | 5,76 |
| 2,4 | 13 | 5,76 | 31,2 | 12,297 | 1,697809 | 0,36 |
| 4,9 | 19 | 24,01 | 93,1 | 18,947 | 28,59041 | 29,16 |
| 4,2 | 18 | 17,64 | 75,60 | 17,085 | 12,14523 | 19,36 |
| 3,0 | 12 | 9,00 | 36,00 | 13,893 | 0,085849 | 2,56 |
| 1,3 | 11 | 1,69 | 14,30 | 9,371 | 17,88444 | 6,76 |
| 1,0 | 8 | 1,00 | 8,00 | 8,573 | 25,27073 | 31,36 |
| 3,0 | 14 | 9,00 | 42,00 | 13,893 | 0,085849 | 0,16 |
| 1,5 | 9 | 2,25 | 13,50 | 9,903 | 13,66781 | 21,16 |
| 4,1 | 16 | 16,81 | 65,60 | 16,819 | 10,36196 | 5,76 |
| 28,9 | 136 | 99,41 | 435,30 | – | 112,4242 | 122,4 |

.

Обчислимо коефіцієнти лінійної регресії за формулами:





Таким чином, шукана регресійна залежність має вигляд:



Нахил лінії регресії 2,66 хвилин на милю – це кількість хвилин, що припадає на одну милю відстані. Координата точки перетину прямої з віссю Y 5,913 хвилин – це час, який не залежить від пройденої відстані, а обумовлюється усіма іншими можливими факторами, явно не врахованими при аналізі.

Обчислимо коефіцієнт детермінації:

 або 91,8%.

Проведемо регресійний аналіз з використанням режиму Регресія MS Excel. Значення параметрів, встановлених в однойменному діалоговому вікні, представлені на рис.2.



Рис. 2.

Сгенерируются результати регресійної статистики, представлені в таблиці 3.

Таблиця 3.

|  |  |
| --- | --- |
| ВИВЕДЕННЯ ПІДСУМКІВ |  |
|  |  |
| *Регресійна статистика* |
| Множинний R | 0,958275757 |
| R-квадрат | 0,918292427 |
| Нормований R-квадрат | 0,90807898 |
| Стандартна помилка | 1,11809028 |
| Спостереження | 10 |

Розглянемо представлену в таблиці 3 регресійну статистику. Величина R-квадрат, звана також мірою визначеності, характеризує якість отриманої регресійної прямої. Ця якість виражається ступенем відповідності між вихідними даними і регресійної моделлю (за розрахунковими даними). Міра визначеності завжди знаходиться в межах інтервалу [0;1]. У нашому прикладі міра визначеності дорівнює 0,91829, що говорить про дуже гарну підгонці регресійної прямої до вихідним даним і збігається з коефіцієнтом детермінації R2, обчисленим за формулою.

Таким чином, лінійна модель пояснює 91,8% варіації часу доставки, що означає правильність вибору фактора (відстані). Не пояснюється  варіації часу поїздки, які обумовлені іншими факторами, що впливають на час поставки, але не включеними в лінійну модель регресії.

Розрахований рівень значущості αр=1,26E-05<0,05 (показник значимість F в таблиці Дисперсійний аналіз) підтверджує значимість R2.

Множинний R - коефіцієнт множинної кореляції R - виражає ступінь залежності незалежних змінних (X) і залежною змінною (Y) і дорівнює квадратному кореню з коефіцієнта детермінації, ця величина приймає значення в інтервалі від нуля до одиниці. У простому лінійному регресійному аналізі множинний R дорівнює коефіцієнту кореляції Пірсона. Дійсно, множинний R в нашому випадку дорівнює коефіцієнту кореляції Пірсона (0,95827), який обчислюється за формулою:

Тепер розглянемо середню частину розрахунків, представлену в таблиці 4 (наведена у скороченому варіанті). Тут дано коефіцієнт регресії  (2,65970168) і зсув по осі ординат, тобто константа  (5,913462144).

Таблиця 4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Коефіцієнти* | *Стандартна помилка* | *t-статистика* | *P-Значення* |
| Y-перетин | 5,913462144 | 0,884389599 | 6,686489927 | 0,00015485 |
| Змінна X 1 | 2,65970168 | 0,280497238 | 9,482095791 | 1,26072E-05 |

Виходячи з розрахунків, можемо записати рівняння регресії таким чином:

 (\*)

Бачимо, що це рівняння, збігається з рівнянням, отриманим нами при розрахунку по МНК вручну з точністю до помилки округлення.

Напрямок зв'язку між змінними визначається на підставі знаків (додатній або від’ємний) коефіцієнта регресії (коефіцієнта ). У нашому випадку знак коефіцієнта регресії позитивний, отже, зв'язок також є позитивною.

Далі перевіримо значимість коефіцієнтів регресії:  і . Порівнюючи попарно значення стовпців *Коефіцієнти* і *Стандартна помилка* в таблиці 4, бачимо, що абсолютні значення коефіцієнтів більше, ніж їх стандартні помилки. До того ж ці коефіцієнти є значущими, про що можна судити за значеннями показника Р-значення в таблиці 4, які менше заданого рівня значущості α=0,05.

Таблиця 5

|  |  |  |
| --- | --- | --- |
| ВИВЕДЕННЯ ЗАЛИШКУ |  |  |
|  |  |  |  |
| *Состереження* | *Передбачення Y* | *Залишки* | *Стандартні залишки* |
| 1 | 15,22241803 | 0,777581975 | 0,737641894 |
| 2 | 12,29674618 | 0,703253823 | 0,667131568 |
| 3 | 18,94600038 | 0,053999622 | 0,051225961 |
| 4 | 17,0842092 | 0,915790799 | 0,868751695 |
| 5 | 13,89256718 | -1,892567185 | -1,795356486 |
| 6 | 9,371074328 | 1,628925672 | 1,545256778 |
| 7 | 8,573163824 | -0,573163824 | -0,543723571 |
| 8 | 13,89256718 | 0,107432815 | 0,101914586 |
| 9 | 9,903014664 | -0,903014664 | -0,8566318 |
| 10 | 16,81823903 | -0,818239033 | -0,776210624 |

У таблиці 5. представлені результати виведення залишків. За допомогою цієї частини звіту ми можемо бачити відхилення кожної точки від побудованій лінії регресії. Найбільше абсолютне значення залишку в нашому випадку - 1,89256, найменше - 0,05399. Для кращої інтерпретації цих даних скористаємося графіком вихідних даних і побудованої лінією регресії, представленими на рис. 3. Як бачимо, лінія регресії добре "підігнана" під значення вихідних даних.



Рис.3.

Приблизними, але самим простим і наочним способом перевірки задовільності регресійної моделі є графічне представлення відхилень.

Рис. 4. Графік відхилень

Відкладемо відхилення  по осі Y, для кожного значення . Якщо регресійна модель близька до реальної залежності, то відхилення будуть носити випадковий характер і їх

Рис.5

сума буде близька до нуля. У розглянутому прикладі .

Зазвичай мірою помилки регресійної моделі служить середнє квадратичне відхилення



Для нормально розподілених процесів приблизно 67% точок знаходиться в межах одного відхилення від лінії регресії і 95% - у межах (на малюнку 8 трубки А і В відповідно).

Вирішимо задачу прогнозування. Оскільки коефіцієнт детермінації R2 має досить високе значення і відстань 2 милі, для якого треба зробити прогноз, перебуває в межах діапазону вихідних даних (таблиця 1), то ми можемо використовувати отримане рівняння лінійної регресії для прогнозування

 хвилин.

При прогнозах на відстані, що не входять в діапазон вихідних даних, не можна гарантувати справедливість отриманої моделі. Це пояснюється тим, що зв'язок між часом і відстанню може змінюватися по мірі збільшення відстані. На час далеких перевезень можуть впливати нові фактори такі, як використання швидкісних шосе, зупинки на відпочинок, обід і т. п.

**Лабораторна работа 2**

Завдання. По території регіону наводяться дані за 199X р.

Таблиця 1

|  |  |  |
| --- | --- | --- |
| Номер регіону | Середньодушовий прожитковий мінімум у день одного працездатного, руб., x | Середньоденна заробітна плата, руб., y |
| 1 | 78 | 133 |
| 2 | 82 | 148 |
| 3 | 87 | 134 |
| 4 | 79 | 154 |
| 5 | 89 | 162 |
| 6 | 106 | 195 |
| 7 | 67 | 139 |
| 8 | 88 | 158 |
| 9 | 73 | 152 |
| 10 | 87 | 162 |
| 11 | 76 | 159 |
| 12 | 115 | 173 |

Вимагається:

1. Побудувати лінійне рівняння парної регресії *y* по *x*.

2. Розрахувати лінійний коефіцієнт парної кореляції, коефіцієнт детермінації і середню помилку апроксимації.

3. Оцінити статистичну значимість рівняння регресії в цілому і окремих параметрів регресії і кореляції з допомогою *F* - критерію Фішера та *t* -критерію Стьюдента.

4. Виконати прогноз заробітної плати *y* при прогнозному значення середньодушового прожиткового мінімуму *x*, становить 107% від середнього рівня.

5. Оцінити точність прогнозу, розрахувавши помилку прогнозу і його довірчий інтервал.

6. На одному графіку відкласти вихідні дані і теоретичну пряму.

Рішення

1. Для розрахунку параметрів рівняння лінійної регресії будуємо розрахункову таблицю 2.

Таблиця 2

 

Знаходимо параметри регресії

;

.

Отримано рівняння регресії:

.

Параметр регресії дозволяє зробити висновок, що з збільшенням середньодушового прожиткового мінімуму на 1 руб. середньоденна заробітна плата зростає в середньому на 92 коп.

Після знаходження рівняння регресії заповнюємо стовпці 7-10 таблиці 2.

2. Тісноту лінійного зв'язку оцінить коефіцієнт кореляції:

.

Оскільки значення коефіцієнта кореляції більше 0,7, то це говорить про

наявність досить тісного лінійного зв'язку між ознаками.

Коефіцієнт детермінації:



Це означає, що 52% варіації заробітної плати (*y*) пояснюється варіацією фактора *x* – середньодушового прожиткового мінімуму.

Якість моделі визначає середня помилка апроксимації:

.

Якість побудованої моделі оцінюється як добрий, так як  не перевищує 10%.

3. Оцінку статистичної значущості рівняння регресії в цілому проведемо з допомогою *F* -критерію Фішера. Фактичне значення *F* - критерію за формулою складе

.

Табличне значення критерію при п'ятивідсотковому рівні значущості і ступенях свободи  і становить . Так як , то рівняння регресії визнається статистично значущим.

Оцінку статистичної значущості параметрів регресії і кореляції здійснимо за допомогою t -статистики Стьюдента і шляхом розрахунку довірчого інтервалу кожного з параметрів.

Табличне значення t -критерію для ступенів свободи  і рівня значущості  складе .

Визначимо стандартні помилки , ,  (залишкова дисперсія на одну ступінь свободи ):

;

;

.

Тоді

;

;
.

Фактичні значення t -статистики перевищують табличне значення:

; ; ,

тому параметри ,  і  не випадково відрізняються від нуля, а статистично значущі.

Розрахуємо довірчі інтервали для параметрів регресії  і . Для цього визначимо граничну помилку для кожного показника:

;

.

Довірчі інтервали

 і ;

 і .

Аналіз верхньої і нижньої меж довірчих інтервалів приводить до висновку про те, що з імовірністю  параметри  і , перебуваючи у зазначених межах, не приймають нульових значень, тобто є статистично значущими і істотно відмінні від нуля.

4. Отримані оцінки рівняння регресії дозволяють використовувати його для прогнозу. Якщо прогнозне значення прожиткового мінімуму складе:  руб., тоді індивідуальне прогнозне значення заробітної плати складе:  руб.

5. Помилка прогнозу становитиме:

.

Гранична помилка прогнозу, яка в 95% випадків не буде перевищена, складе:

.

Довірчий інтервал прогнозу:

 і .

Виконаний прогноз середньомісячної заробітної плати є надійним

( ) і знаходиться в межах від 131,92 руб. до 190,66 руб.

6. На закінчення рішення задачі побудуємо на одному графіку вихідні дані і теоретичну пряму (рис. 1):



Рис. 1.

**Рішення типової задачі в MS Excel**

C допомогою інструмента аналізу даних Регресія можна отримати результати регресійної статистики, дисперсійного аналізу, довірчих інтервалів, залишки і графіки підбору лінії регресії.

Якщо в меню сервіс ще немає команди **Аналіз даних**, то необхідно зробити наступне. В головному меню послідовно вибираємо **Сервіс→Надбудови** і встановлюємо «прапорець» в рядку **Пакет аналізу** (рис. 1)



Рис. 1

Далі слідуємо за таким планом.

1. Якщо вихідні дані вже внесені, то вибираємо **Сервіс→Аналіз→Регресія**.

2. Заповнюємо діалогове вікно введення даних і параметрів виводу (рис. 2)



Рис. 2

Тут:

*Вхідний інтервал Y* – діапазон, що містить дані результативної ознаки;

*Вхідний інтервал X* – діапазон, що містить дані признакафактора;

*Мітки* – «прапорець», який вказує, утримуй перший рядок назви стовпців;

*Константа – нуль* – «прапорець», що вказує на наявність або відсутність вільного члена в рівнянні;

*Вихідний інтервал* – достатньо вказати верхню ліву клітинку майбутнього діапазону;

*Новий робочий лист* – можна вказати довільне ім'я нового листа (або не вказувати, тоді результати виводяться на знову створений аркуш).

Отримуємо наступні результати для розглянутого вище приклад:



Рис. 3

Звідки виписуємо, округляючи до 4 знаків після коми і переходячи

до наших позначень:

Рівняння регресії:

.

Коефіцієнт кореляції:

.

Коефіцієнт детермінації:

.

Фактичне значення F -критерію Фішера:

.

Залишкова дисперсія на одну ступінь свободи:

.

Корінь квадратний із залишкової дисперсії (стандартна помилка):

.

Стандартні помилки параметрів регресії:

, .

Фактичні значення t -критерію Стьюдента:

, .

Довірчі інтервали:

,

.

Як бачимо, знайдені всі розглянуті вище параметри і характеристики рівняння регресії, за винятком середньої помилки апроксимації (значення *t* -критерію Стьюдента для коефіцієнта кореляції збігається з ). Результати «ручного рахунку» від машинного відрізняються незначно (відмінності пов'язані з помилками округлення).

**Лабораторна робота 3**

**Підготовка до роботи**

1. Ознайомитися з теоретичним матеріалом про парну лінійну регресію.
2. Повторити основні відомості про функціональні можливості пакета MS EXCEL.

**Завдання до роботи**

1. Для парної лінійної регресії на основі статистичних даних показника *y* (табл.1.2) і фактора *x* (табл.1.1) знайти оцінки:
* параметрів лінії регресії *, *;
* коефіцієнта кореляції ;
1. Використовуючи критерій Фішера, з надійністю =0,05 оцінити адекватність прийнятої моделі статистичним даним.
2. Якщо модель адекватна статистичним даним, то знайти:
* з надійністю =0,05 надійні зони базисних даних;
* прогноз показника та його надійні інтервали;
* коефіцієнт еластичності для базисних даних і прогнозу.
1. Побудувати графіки:
* статистичних даних;
* лінії регресії і її довірчої зони;
* коефіцієнта еластичності.

Таблиця 1

**Значення фактора**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | *x1* | *x2* | *x3* | *x4* | *x5* | *x6* | *x7* | *x8* | *x9* | *x10* | *x11* |
| 1 | 2,06 | 2,53 | 2,17 | 3,65 | 3,22 | 2,16 | 4,57 | 2,25 | 6,15 | 1,86 | 2,07 |
| 2 | 2,58 | 3,54 | 2,90 | 3,82 | 3,87 | 2,65 | 5,42 | 2,98 | 5,66 | 1,91 | 3,22 |
| 3 | 3,14 | 3,84 | 3,29 | 3,76 | 4,95 | 3,49 | 5,29 | 2,15 | 7,50 | 2,14 | 3,04 |
| 4 | 3,54 | 3,84 | 4,13 | 5,24 | 5,10 | 3,16 | 6,33 | 2,71 | 6,90 | 3,39 | 3,42 |
| 5 | 4,18 | 4,22 | 5,25 | 5,03 | 5,98 | 3,85 | 7,63 | 3,70 | 8,31 | 3,95 | 5,23 |
| 6 | 4,78 | 4,81 | 4,92 | 5,52 | 7,28 | 4,58 | 7,53 | 4,59 | 8,25 | 4,3 | 5,7 |
| 7 | 5,11 | 6,53 | 5,79 | 5,62 | 6,90 | 5,33 | 7,73 | 4,77 | 9,39 | 5,1 | 6,53 |
| 8 | 5,67 | 5,82 | 5,87 | 6,98 | 7,54 | 5,89 | 8,44 | 5,34 | 9,73 | 5,47 | 6,41 |
| 9 | 6,02 | 6,43 | 6,99 | 6,91 | 7,91 | 6,20 | 9,49 | 5,45 | 9,33 | 5,97 | 6,68 |
| 10 | 6,65 | 7,73 | 7,04 | 7,95 | 8,40 | 6,39 | 9,18 | 6,00 | 10,5 | 6,16 | 7,46 |
| 11 | 7,05 | 8,19 | 8,14 | 7,24 | 8,14 | 6,95 | 10,14 | 6,25 | 11,1 | 6,46 | 6,83 |
| 12 | 7,52 | 7,65 | 8,06 | 9,27 | 8,76 | 7,25 | 9,94 | 6,79 | 11,51 | 6,07 | 6,34 |
| 13 | 8,03 | 9,31 | 8,57 | 8,46 | 9,67 | 7,80 | 10,92 | 8,24 | 12,42 | 6,71 | 8,19 |
| 14 | 8,56 | 9,26 | 9,45 | 10,3 | 10,28 | 8,47 | 11,89 | 8,51 | 12,4 | 7,16 | 7,19 |
| 15 | 9,03 | 9,86 | 9,06 | 10,72 | 10,59 | 9,22 | 11,14 | 9,15 | 13,14 | 8,81 | 9,72 |
| *хр* | 9,52 | 9,69 | 10,3 | 10,05 | 11,58 | 9,32 | 11,73 | 9,78 | 12,56 | 8,07 | 8,71 |

Таблиця 2

**Значення показника**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| N | *y1* | *y2* | *y3* | *y4* | *y5* | *y6* | *y7* | *y8* | *y9* | *y10* | *y11* |
| 1 | 7,24 | 10,89 | 16,21 | 12,11 | 15,21 | 16,62 | 10,22 | 12,50 | 19,66 | 14,87 | 22,68 |
| 2 | 8,02 | 11,92 | 17,75 | 12,3 | 15,42 | 17,63 | 10,58 | 13,88 | 20,53 | 15,78 | 23,89 |
| 3 | 9,28 | 12,45 | 18,39 | 13,82 | 16,44 | 19,22 | 12,01 | 15,16 | 21,31 | 16,79 | 24,32 |
| 4 | 10,12 | 13,27 | 18,87 | 14,84 | 17,93 | 19,36 | 12,84 | 16,06 | 22,59 | 18,03 | 25,97 |
| 5 | 11,12 | 14,12 | 19,6 | 15,86 | 18,52 | 20,52 | 13,28 | 16,66 | 23,27 | 18,29 | 26,23 |
| 6 | 12,19 | 15,23 | 21,21 | 16,41 | 19,8 | 21,95 | 15,13 | 17,65 | 24,44 | 19,93 | 27,60 |
| 7 | 13,01 | 16,07 | 21,84 | 17,8 | 20,76 | 22,45 | 15,84 | 18,46 | 25,85 | 20,32 | 28,13 |
| 8 | 14,12 | 17,4 | 23,00 | 18,61 | 21,3 | 23,56 | 17,08 | 19,54 | 26,74 | 21,18 | 29,84 |
| 9 | 15,21 | 18,68 | 24,44 | 19,57 | 22,25 | 24,9 | 17,99 | 20,58 | 27,36 | 22,47 | 30,31 |
| 10 | 16,29 | 19,46 | 25,36 | 21,26 | 24,14 | 25,53 | 18,32 | 21,77 | 28,37 | 23,47 | 31,52 |
| 11 | 17,01 | 20,52 | 25,54 | 21,08 | 24,17 | 26,11 | 19,49 | 22,15 | 29,22 | 24,07 | 32,27 |
| 12 | 18,03 | 21,32 | 27,14 | 22,99 | 25,66 | 28,02 | 20,59 | 23,80 | 30,50 | 25,57 | 33,77 |
| 13 | 19,19 | 22,58 | 27,95 | 23,43 | 26,5 | 28,37 | 21,35 | 24,79 | 31,21 | 27,07 | 34,66 |
| 14 | 20,21 | 23,73 | 28,99 | 24,63 | 27,46 | 29,46 | 23,20 | 25,57 | 32,56 | 27,62 | 35,93 |
| 15 | 21,22 | 25,02 | 30,8 | 25,41 | 29,02 | 30,42 | 24,21 | 27,18 | 33,66 | 28,42 | 36,97 |

**Порядок виконання роботи**

* 1. Запустіть програму EXCEL. Згідно з номером варіанту, виданого викладачем, занесемо початкові значення фактора *x* в блок **А3:А17**, а значення показника *y* в блок **В3:В17**. Надалі, при формуванні розрахункової таблиці рекомендується дотримуватися форми, наведеної в табл.1.3. В першому рядку необхідно написати назву лабораторної роботи, а другий рядок буде використаний для занесення даних, які будуть отримуватися під час розрахунків. В табл.1.3 затемнення комірки означає, що в ній міститься формула, або конкретне числове значення якогось параметра.
	2. Для обчислення параметрів ** та ** потрібно підрахувати добутки  та  і відповідні суми:
* для знаходження добутку  використаємо блок **С3:С17**. Для клітинки **С3** формула запишеться: =А3\*В3, дану формулу копіюємо в блок **С4:С17**;
* значення  занесемо в блок **D3:D17**. Для клітинки **D3** формула запишеться: =А3^2, дану формулу копіюємо в блок **D4:D17**.
	1. Для підрахунку сум будемо використовувати математичну функцію **СУММ(блок)**:
* в 20 рядку будемо знаходити суми відповідних значень. Значення  підрахуємо в **А20** за формулою =СУММ(А3:А17);

# *Таблиця 1.3*

## ***Розрахункова таблиця до лабораторної роботи №1***

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J | K | L |
| 1 | **Парна лінійна регресія** |
| 2 | *xi* | *yi* | *xi yi* | *xi2* |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |
| … |  |  |  |  |  |  |  |  |  |  |  |  |
| 17 |  |  |  |  |  |  |  |  |  |  |  |  |
| 18 | *xр* |  |  |  | р |  |  |  |  |  |  |  |
| 19 |  |  |  |  |  |  |  |  |  |  |  |  |
| 20 |  |  |  |  |  |  |  |  |  |  |  |  |
| 21 | *n=* |  | = |  |  |  |  |  |  |  |  |  |
| 22 | *b1=* |  | = |  | *t=* |  |  | Fрозр= |  | Fкр= |  |  |
| 23 | *b0=* |  | = |  | *rxy=* |  |  |  |  |  |  |  |

* цю формулу копіюємо в комірки **B20** - , **C20** - , **D20** - .
	1. Кількість значень факторів *n=15* занесемо в комірку **B21**.
	2. Підставимо адреси знайдених сум в формули (1.2) для розрахунку ** та *:*
* формулу для визначення параметра ** занесемо в комірку **В22** =(В21\*C20-A20\*B20)/(B21\*D20-A20^2);
* значення параметра ** занесемо в **B23** =(B20-B22\*A20)/B21.
	1. Для знаходження розрахункових значень показника  використовується рівняння регресії (1.1). Занесемо їх в блок **E3:E17**:
* для **Е3** маємо формулу =B$22\*A3+B$23, копіюємо її в блок **Е4:Е17;**
* в **Е20** підрахуємо суму  за формулою =СУММ(Е3:Е17) або просто копіюємо формулу з **D20**;
* при правильному виконанні розрахунків , тобто значення комірок **В20** та **Е20** повинні співпадати.
	1. Для оцінки коефіцієнта кореляції будемо використовувати статистичну функцію **КОРРЕЛ(блок знач.Х; блок знач.Y)**. Для цього в комірку **F23** заносимо =КОРРЕЛ(А3:А17;В3:В17).
1. Для оцінки адекватності прийнятої моделі використовують формулу (1.2). Необхідно визначити середнє значення для *x* та *y* за допомогою статистичної функції **СРЗНАЧ(блок):**
* середнє  визначимо в **D21** =CРЗНАЧ(А3:А17);
* середнє  - в **D22** =CРЗНАЧ(В3:В17).
1. Значення  обчислюємо в **F3:F17**. Для комірки **F3** формула =(B3-E3)^2, яку копіюємо в блок **F4:F17**.
2. Аналогічно значення  обчислюємо в **G3:G17**. Для комірки **G3** формула =(Е3-D$22)^2, яку копіюємо в блок **G4:G17**.
3. Cуми цих величин знайдемо в 20 рядку, в **F20** і  в **G20**.
4. Знайдемо розрахункове значення Фішера *F* в **І22** за формулою =(G20\*(B21-2))/(F20).
5. Табличне значення *Fкр* знаходимо за допомогою статистичної функції **FРАСПОБР(****;1;n-2)** в **К22** за формулою =FРАСПОБР(0,05;1;13).
	1. Для визначення надійної зони базисних даних необхідно обчислити:
* значення функції Ст’юдента  за допомогою статистичної функції **СТЬЮДРАСПОБР(;k)**. Підрахуємо його в **F22** =СТЬЮДРАСПОБР(0,05;14);
* оцінку  в **D23** =КОРЕНЬ(F20/(B21-2));
* значення  в **Н3:Н17**, занесемо в **Н3** формулу =(А3-D$21)^2 і скопіюємо її в блок **Н4:Н17**;
* підрахуємо суму  в **Н20**;
* підставляємо отримані адреси в формулу і знаходимо  в **І3:І17**. В **І3** заносимо формулу =F$22\*D$23\*КОРЕНЬ(1/B$21+H3/H$20).
* значення  знаходимо за формулами (1.7) в **J3:L17**, **J3** =E3-I3. Аналогічно  в **K3:K17**, **K3** =E3+I3.

3.2. Згідно з виразом (1.8) визначаємо коефіцієнт еластичності для 15-ти значень *x* в **L3:L17** за формулою **L3** =B$22\*A3/E3.

3.3. Прогнозне значення фактора *xp* (табл.1.1) занесемо в 18 рядок, в комірку **А18**.

3.4. Значення прогнозу показника підрахуємо в **Е18** за формулою =В22\*А18+В23, або просто копіюємо формулу з **Е17** в **Е18**.

3.5. Для прогнозних значень визначимо: значення  в **Н18**,  в **І18**,  в **J18**,  в **K18**,  в **L18**. Для цього достатньо просто скопіювати формули з вище розташованих відповідних комірок.

4. Для наочного уявлення результатів будуємо графіки:

* статистичних даних (залежність показника *y* (В3:В17) від фактора *x* (A3:А17)) і лінії регресії (залежність значень  (E3:E17) від фактора *x* (A3:А17));
* лінії регресії (дублюємо) та її довірчої зони ,  (відповідно залежність значень (J3:J17) та (K3:K17) від фактора *x* (A3:А17)).
* гістограма для коефіцієнта еластичності  (L3:L17).

Приклад графіків, що ілюструють результати даної лабораторної роботи, наведено в додатку А.

Орієнтовний порядок побудови графіків може бути наступним:

4.1. Наводимо курсор на ***Мастер диаграмм***і натискаємо ліву клавішу миші.

4.2. Як тільки відпустимо ліву кнопку миші, відкривається перше діалогове

вікно ***Тип диаграммы***.На вкладці ***Стандартные*** вибираємо тип діаграми,

наприклад, ***Точечная*** і її вид – **5**. Для переходу до наступного діалогового

вікна натискаємо кнопку ***Далее>.***

4.3. У вікні ***Источник данных диаграммы*** на вкладці ***Диапазон данных*** вказуємо, що дані знаходяться в стовпцях, а на вкладці ***Ряд*** в однойменному полі, шляхом натискання на кнопку ***Добавить*,** створюємо необхідну кількість рядів, які будуть відображені на графіку. Для роботи з відповідним рядом необхідно його відмітити мишою. Для кожного ряду в полі ***Имя*** вводиться ім’я, а в полях ***Значения*** ***Х*** та ***Значения*** ***Y*** – відповідно діапазони значень фактора та показника. Для цього курсор ставиться у відповідному полі, наприклад, ***Значения*** ***Х*** після чого натискається кнопка, яка знаходиться в кінці цього рядка. В результаті програма повертається до роботи з таблицею і користувач має змогу відмітити мишею потрібні для побудови графіка блоки, наприклад: *АЗ-А18.* Або, можна ввести ці ж дані вручну. Для переходу до наступного діалогового вікна натискаємо кнопку ***Далее>.***

4.4. У вікні ***Параметры диаграммы*** на вкладці ***Заголовки*** вводять назву графіка і назву координатнихвісей**.** При необхідності, або ж за бажанням користувача, на інших вкладках змінюється решта параметрів діаграми після чого знову таки натискаємо кнопку ***Далее>***.

4.5. У діалоговому вікні ***Размещение диаграмм*** необхідновказати куди саме ви бажаєте помістити діаграму: на окремий аркуш чи в поле робочого аркуша, наприклад, нижче таблиці. Після останнього вікна натискаємо клавішу ***Готово*,** після чого на робочому аркуші з'являється графік.

Якщо на будь-якому кроці отриманий результат не задовольняє, можна повернутися до попереднього діалогу, натиснувши кнопку ***<Назад.*** Для редагування графіка (або його частин) необхідно навести на нього курсор і натиснути 2 рази на ліву клавішу миші.

5. Перед тим як ***роздрукувати робочий аркуш,*** необхідно в меню ***Файл*** знайти команду ***Параметри страницы,*** після чого відкриється вікно діалогу, яке має такі вкладки: ***Страница, Поля, Колонтитулы, Лист.***

5.1. Вибравши вкладку ***Страница****,* у групі ***Ориентация*** задаємо розташування паперового аркуша при роздрукуванні. За допомогою групи ***Масштаб*** можна зменшити (збільшити) зображення таблиць. При цьому, вибравши опції ***Разместить лист на****,* можна задати, на скількох сторінка паперу розташувати робочий аркуш Excel в довжину та в ширину. Вибір опції ***Установить*** дозволяє задати коефіцієнт масштабування у процентах від нормального розміру таблиці.

5.2. За допомогою вкладки ***Поля*** необхідно встановити відступи від країв паперу та вирівнювання інформації на аркуші. В групі ***Центрирование*** знаходяться перемикачі горизонтальні та вертикальні, які потрібні для визначення положення таблиці на паперовому аркуші по відношенню до його країв.

5.3. За допомогою вкладки ***Лист*** задаємо опції друку для таблиці. У полі ***Выводить на печать диапазон*** вказуємо діапазон комірок, необхідних для друку. Якщо це поле пусте, тоді друкується вся таблиця. При установці перемикача ***Сетка*** в групі ***Печатать*** надрукована таблиця буде мати сітку таку ж, як на екрані, що дуже зручно для зображення таблиць на папері.

5.4. Для друку таблиці вибираємо команду меню ***Файл Печать****.* Тут відмічаємо необхідні умови для друку і натискуємо клавішу ***Enter*** або ***Ok****.*