# Лабораторна робота №8

## Прогнозування на основі лінії тренду з використанням MS Excel: Використання функцій прогресії для розрахунку прогнозів

***Мета:*** Навчитись використовувати засоби *MS Excel* для побудови кількісних прогнозів на наступний часовий період на основі даних за минулий часовий період.

***Література***

Використання надбудови ''Пакет аналізу'' для виконання аналізу складних даних:

[https://support.microsoft.com/uk-ua/office/використання-надбудови-пакет-аналізу-длявиконання-аналізу-складних-даних-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6](https://support.microsoft.com/uk-ua/office/використання-надбудови-пакет-аналізу-для-виконання-аналізу-складних-даних-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6)

Надбудова "Аналіз даних": [https://www.ateasyday.com/articles/programs/nadstrojkaanaliz-dannih-v-eksele.html](https://www.ateasyday.com/articles/programs/nadstrojka-analiz-dannih-v-eksele.html)

## Зміст роботи

### ***Завдання 1. З***а даними спостережень, представлених у вигляді таблиці значень x і y, виконати у середовищі табличного процесора MS Excel:

1. знайти коефіцієнти b0 і b1 рівняння лінійної регресії y = b0 + b1x;
2. оцінити відповідність отриманого рівняння експериментальним

значенням за допомогою коефіцієнта достовірності апроксимації (R^2);

1. виконати прогноз по значенню y для заданого значення x,

використовуючи отримане рівняння лінійної регресії;

1. побудувати аппроксимирующий поліном другого порядку за

даними таблиці;

1. оцінити відповідність отриманого рівняння експериментальним

значенням за допомогою коефіцієнта достовірності апроксимації (R^2);

1. виконати прогноз для заданого значення x, використовуючи

отриманий поліном;

1. порівняти результати прогнозу для рівняння регресії і поліноміальної апроксимації.

Дані для розрахунків вибираються з таблиці відповідно до номеру

варіанта

Варіанти завдань

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***№В*** | ***хі*** | ***2*** | ***4*** | ***5*** | ***6*** | ***8*** | ***9*** | ***11*** | ***12*** | ***13*** | ***15*** | ***16*** | ***18*** | ***Значення х для прогнозу*** |
| 1 | yi | 3 | 6 | 6 | 6 | 10 | 9 | 12 | 14 | 15 | 16 | 17 | 19 | 21 |
| 2 | yi | 5 | 9 | 10 | 13 | 17 | 19 | 23 | 26 | 28 | 28 | 29 | 30 | 20 |
| 3 | yi | 7 | 13 | 16 | 19 | 25 | 28 | 34 | 38 | 41 | 45 | 47 | 50 | 22 |
| 4 | yi | 9 | 15 | 20 | 22 | 33 | 37 | 45 | 50 | 54 | 56 | 60 | 63 | 25 |
| 5 | yi | 11 | 18 | 20 | 31 | 40 | 44 | 56 | 62 | 67 | 70 | 76 | 78 | 24 |
| 6 | yi | 11 | 26 | 31 | 40 | 49 | 58 | 68 | 74 | 80 | 85 | 89 | 93 | 23 |
| 7 | yi | 15 | 29 | 36 | 43 | 57 | 64 | 78 | 86 | 92 | 98 | 103 | 106 | 20 |
| 8 | yi | 17 | 33 | 40 | 48 | 65 | 73 | 88 | 104 | 106 | 114 | 116 | 122 | 21 |
| 9 | yi | 6 | 10 | 12 | 17 | 18 | 20 | 22 | 28 | 29 | 30 | 33 | 35 | 22 |
| 10 | yi | 8 | 14 | 17 | 20 | 26 | 29 | 36 | 39 | 42 | 46 | 49 | 55 | 23 |
| 11 | yi | 9 | 18 | 20 | 22 | 26 | 38 | 47 | 50 | 52 | 54 | 64 | 75 | 24 |
| 12 | yi | 10 | 22 | 32 | 30 | 35 | 48 | 58 | 60 | 70 | 72 | 79 | 80 | 25 |
| 13 | yi | 14 | 20 | 30 | 40 | 42 | 50 | 65 | 70 | 78 | 88 | 100 | 110 | 21 |
| 14 | yi | 15 | 30 | 37 | 44 | 58 | 66 | 80 | 87 | 94 | 98 | 105 | 110 | 22 |
| 15 | yi | 18 | 34 | 43 | 51 | 67 | 75 | 92 | 99 | 107 | 110 | 118 | 125 | 23 |

## Методичні рекомендації

MS Excel має засоби кількісного прогнозування, які дають змогу зробити прогноз шляхом поширення (екстраполяції) даних на наступний часовий період на основі даних за минулий часовий період.

*Часовий ряд* - числова послідовність даних спостережень, що характеризують зміну певної величини, наприклад, певного економічного показника в часі. Кожен елемент часового ряду називають рівнем ряду, він відповідає певному моментові часу.

У MS Excel лінію рівняння регресії називають лінією тренду. Вона вказує тенденцію зміни даних, її застосовують для складання прогнозів.

Лінію тренду будують на основі діаграми. Щоб побудувати лінію тренду, можна використовувати один із п'яти типів апроксимації:

* лінійна;
* логарифмічна;
* степенева;
* експоненціальна;
* поліноміальна.

### Приклад виконання завдання.

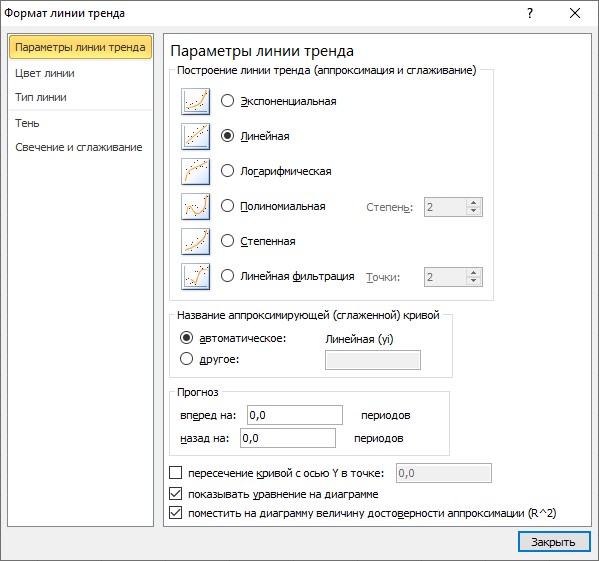
Виконаємо завдання для даних, представлених в таблиці нижче:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***хі*** | ***1*** | ***3*** | ***4*** | ***5*** | ***6*** | ***8*** | ***10*** | ***11*** | ***12*** | ***13*** | ***14*** |
| yi | 11 | 21 | 27 | 45 | 55 | 68 | 78 | 83 | 86 | 89 | 92 |

Для побудови прогнозу будемо використовувати x = 16.

За даними таблиці побудуємо діаграму точкового типу.

До діаграми додаємо лінію тренду , тип лінії тренду *Лінійна*, а на вкладці *Параметри* встановимо прапорці показувати рівняння на діаграмі і помістити на діаграму величину достовірності апроксимації (R^2). Лінія тренду, відповідна рівнянню лінійної регресії, побудованої за методом найменших квадратів буде додана до діаграми. Крім того, буде виведено рівняння регресії і коефіцієнт достовірності апроксимації (рис.1).



ПКМ на графіку

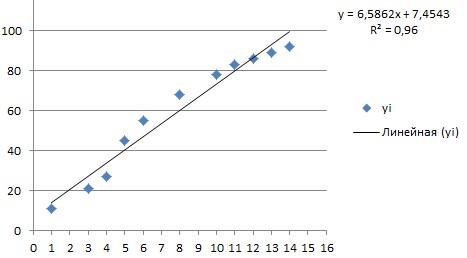


Рис.1. Діаграма і Лінійна лінія тренду

Рівняння регресії матиме вигляд y= 6,5862x + 7,4543, оцінка достовірності апроксимації R2=0,96. Визначимо прогноз для x = 16, виконавши обчислення за формулою, отримаємо y = 112, 834

На тій же діаграмі побудуємо ще одну лінію тренду, вибравши тип *Поліноміальна* степені 2 і встановивши параметри для виведення рівняння і величини достовірності апроксимації (рис.2). На цей раз рівняння буде мати наступний вигляд y= -0,3077x2+11,335x-5,4593, оцінка достовірності апроксимації R2=0,984. Визначимо прогноз по y для x = 16, виконавши обчислення отримаємо y = 97,1295.

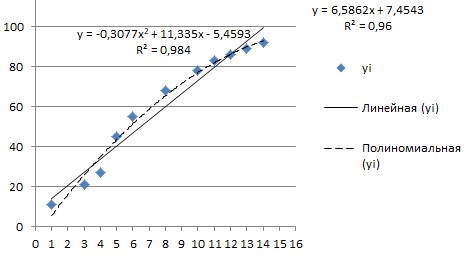


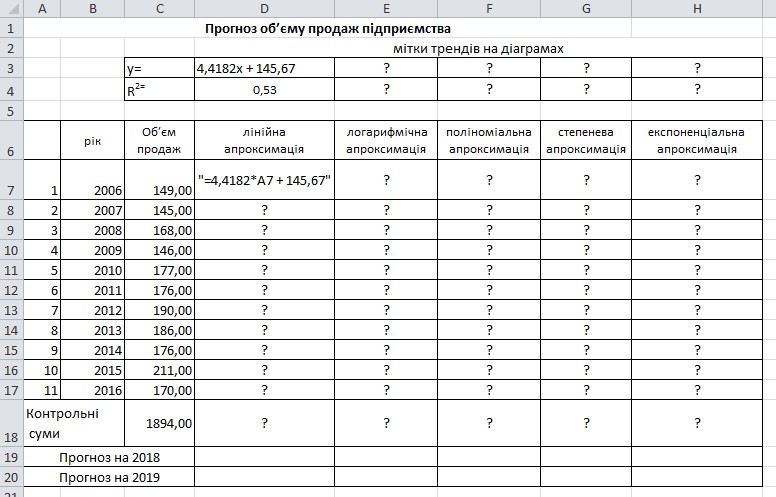
Рис.2. Діаграма і лінії тренду Лінійна і Поліноміальна

Аналіз результатів дозволяє припустити, що прогноз з використанням поліноміальної апроксимації є більш точним, тому що ступінь достовірності апроксимації в цьому випадку вище.

### ***Завдання 2.*** Є дві величини x і y, наприклад, обсяг реалізованої продукції за ряд років. Необхідно з'ясувати яка з найбільш поширених функціональних залежностей підходить для опису процесу реалізації товару, і який результат за обсягами продажів можна очікувати в наступні роки роботи фірми.

Для того, щоб побудувати прогноз розвитку будь-якої ситуації на практиці часто необхідно знати закономірність зміни досліджуваної величини або об'єкта.

Для виявлення тенденцій розвитку процесу продажів необхідно побудувати тренди і здійснити їх аналіз. Побудувати і проаналізувати, як описують процес динаміки продажів лінійна, логарифмічна, поліноміальна, статечна і експоненціальна залежності.



### ***Завдання 3.*** Застосувати вбудовані функції прогнозування для складання кількісного прогнозу на 2023 та 2024 роки, для прийняття рішення стосовно плану випуску продукції на найближчі два роки, відповідно статистичних даних представлених у наступній таблиці:

Варіанти завдань

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Обсяг продаж (тис. од.) | | | |  |  |  |  |
|  | ***рік*** | ***2017*** | ***2018*** | ***2019*** | ***2020*** | ***2021*** | ***2022*** | ***2023*** | ***2024*** |
| ***1*** | ***обсяг*** | 230 | 290 | 305 | 320 | 338 | 356 | ? | ? |
| ***2*** | ***обсяг*** | 150 | 218 | 299 | 305 | 325 | 405 | ? | ? |
| ***3*** | ***обсяг*** | 99 | 112 | 204 | 297 | 315 | 395 | ? | ? |
| ***4*** | ***обсяг*** | 258 | 306 | 455 | 325 | 486 | 499 | ? | ? |
| ***5*** | ***обсяг*** | 112 | 158 | 254 | 297 | 315 | 382 | ? | ? |

## Методичні рекомендації

MS Excel має функції прогнозування, що належать до категорії статистичних функцій:

 **Функція ПРЕДСКАЗ.** Ця функція обчислює одне значення рівняння лінійної регресії, і має такий синтаксис:

# ПРЕДСКАЗ(х; відомі\_знач\_у; відомі\_знач\_х),

де **х –** значення незалежної величини; **відомі\_знач\_у** – масив відомих значень залежної величини, значення якої спостерігаються; **відомі\_знач\_х** – масив відомих значень незалежної величини (наприклад, значення часу), для яких відомі значення залежної величини, яка спостерігається. Розмір масивів **відомі\_знач\_у** та **відомі\_знач\_х** повинен бути однаковим.

Якщо аргумент **відомі\_знач\_х** відсутній, то вважається, що це масив {1; 2; 3; …..; n }, де n – розмір масивів **відомі\_знач\_у** та **відомі\_знач\_х.**

 **Функція ТЕНДЕНЦИЯ.** Ця функція обчислює значення

рівняння лінійної регресії для цілого діапазону значень незалежної змінної як для випадку одновимірного так і для випадку багатовимірного рівняння регресії. Багатовимірна лінійна модель регресії має вигляд. Функція має такий синтаксис:

**ТЕНДЕНЦИЯ (відомі\_знач\_у; відомі\_знач\_х; нові\_знач\_ х; стала**), де **нові\_знач\_ х –** масив значень незалежної величини, що спостерігається (наприклад, певне значення часу), **відомі\_знач\_у** – масив відомих значень залежної величини, значення якої спостерігаються; **відомі\_знач\_х** – масив відомих значень незалежної величини (наприклад, значення часу), **стала** – логічне значення, яке вказує, чи потрібно, щоб стала **b** у формулі дорівнювала нулю: *істина або відсутність цього аргументу* – **b** обчислюється, *хибність* – **b** вважається рівним 0.

Розмір масивів **відомі\_знач\_у** та **відомі\_знач\_х** повин бути однаковим.

Для багатовимірного рівняння регресії потрібно задавати масиви **відомі\_знач\_х** та **нові\_знач\_ х** для кожної незалежної змінної. Якщо аргумент **нові\_знач\_ х** відсутній, то вважається, що масив **нові\_знач\_ х** співпадає з масивом **відомі\_знач\_х.**

Для *експоненціальної апроксимації* статистичних даних , де с, b – сталі, MS Excel має функцію прогнозування **РОСТ.**

**Функція РОСТ.** Має такий синтаксис:

**РОСТ (відомі\_знач\_у; відомі\_знач\_х; нові\_знач\_ х; стала**), де **нові\_знач\_ х –** масив значень незалежної величини, що спостерігається (наприклад, певне значення часу), **відомі\_знач\_у** – масив відомих значень залежної величини, значення якої спостерігаються; **відомі\_знач\_х** – масив відомих значень незалежної величини (наприклад, значення часу), для яких відомі значення залежної величини, яка спостерігається; **стала** – логічне значення, яке вказує, чи потрібно, щоб стала **b** у формулі дорівнювала нулю: *істина або відсутність цього аргументу* – **b** обчислюється, *хибність* – **b** вважається рівним 0.

**відомі\_знач\_х** та **нові\_знач\_ х** для кожної незалежної змінної. Якщо аргумент **нові\_знач\_ х** відсутній, то вважається, що масив **нові\_знач\_ х** співпадає з масивом **відомі\_знач\_х.**

Якщо аргумент **відомі\_знач\_х** відсутній, то вважається, що це масив {1; 2; 3; …..; n }, де n – розмір масивів **відомі\_знач\_у** та **відомі\_знач\_х.**

### ***Завдання 4.*** Скласти прогноз доходу на аукціоні (щотижневому) за допомогою ковзного середнього та графічного аналізу.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Період | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Дохід  (тис. грн.) | 985,9 | 835,6 | 901,4 | 975,6 | 994,4 | 1104,6 | 1256,8 | 1185,6 | 1057,11 | 1011,24 | 969,56 |

## Методичні рекомендації

Ковзна середня−це методика прогнозування, що спрощує аналіз тренду (тенденції), шляхом згладжування флуктуацій результатів вимірів за деякий період часу. Ці флуктуації можуть бути викликані випадковим «шумом», тобто побічним результатом методики виміру. Наприклад, результати виміру росту дитини будуть змінюватися залежно від похибки лінійки та й від того, чи стоїть дитина прямо, чи сутулячись. Флуктуації у вимірах можуть викликатися і іншими умовами. Це вносить систематичну помилку. Наприклад, виторг за місяць може залежати від кількості робочих днів або від того, що провідний продавець перебував у відпустці.

Застосовувати метод ковзної середньої в Excel найкраще за допомогою найпотужнішого інструменту статистичної обробки даних, який називається Пакетом аналізу. Крім того, в цих же цілях можна використовувати вбудовану функцію Excel СРЗНАЧ.

## Контрольні запитання

1. У чому полягає сутність поняття часовий ряд? Наведіть приклади часових рядів.
2. Визначте сутність регресії.
3. Що таке тренд?
4. У чому полягає сутність лінії тренду?
5. Які типи апроксимації застосовуються в *MS Excel* для побудови лінії тренду?
6. Який найвищий ступінь полінома при поліноміальній апроксимації в *MS Excel?*
7. Для яких типів діаграм можна побудувати лінію тренду?
8. Яким чином можна відобразити на діаграмі рівняння лінії тренду та достовірність апроксимації?
9. За якої достовірності апроксимації за допомогою лінії тренду можна прогнозувати дані на майбутнє?
10. Як оцінити, яка лінія тренду краще прогнозуватиме дані?
11. Яким чином за допомогою лінії тренду побудувати прогнозні дані?
12. Які вбудовані функції має *MS Excel* для прогнозування на основі лінійної апроксимації?
13. Які вбудовані функції має *MS Excel* для прогнозування на основі експоненціальної апроксимації?
14. Чим відрізняються функції ПРЕДСКАЗ і ТЕНДЕНЦИЯ?
15. Яка спільна властивість характерна для функцій ТЕНДЕНЦИЯ і РОСТ?
16. Яку відмінність має лінійна фільтрація порівняно з іншими типами апроксимації часового ряду?