



REGIONAL OFFICE FOR

**World Health  
Organization**

---

**Europe**

## Обробка «сирих» даних ПЛР\_частина 2

Лора Чернишова,  
Лабораторний офіцер  
Бюро ВОЗ в Україні,  
[chernyshoval@who.int](mailto:chernyshoval@who.int)

# Серія вебінарів «Обробка даних ПЛР»

**Вебінар 1 – базові поняття.**  
Масштабування, логарифмування, зведення ПЛР-кривих на єдину базову лінію.  
**24.03.2023**



**Вебінар 2 – лінійна апроксимація, математичні моделі обробки ПЛР-кривих.**  
**21.04.2023**



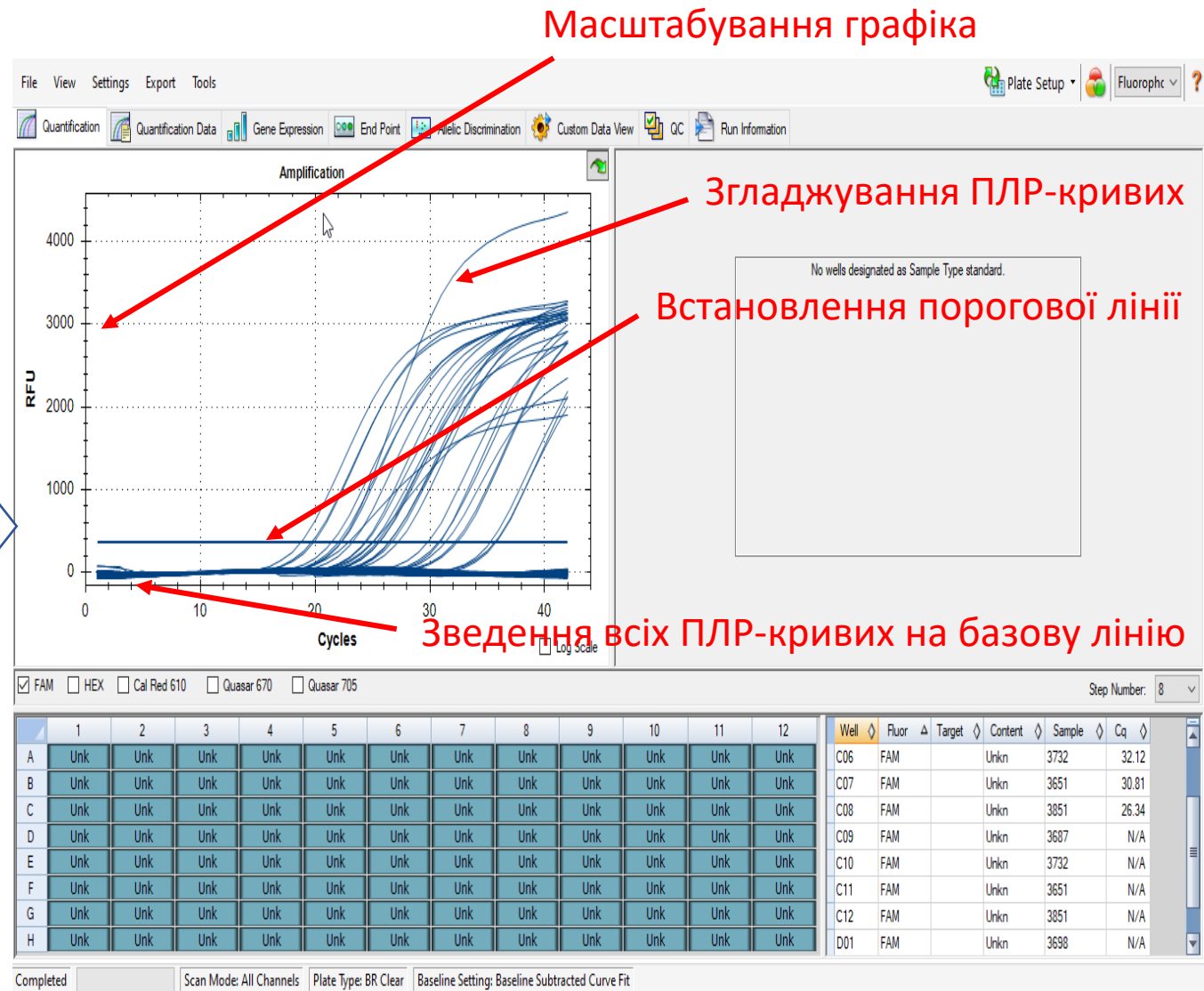
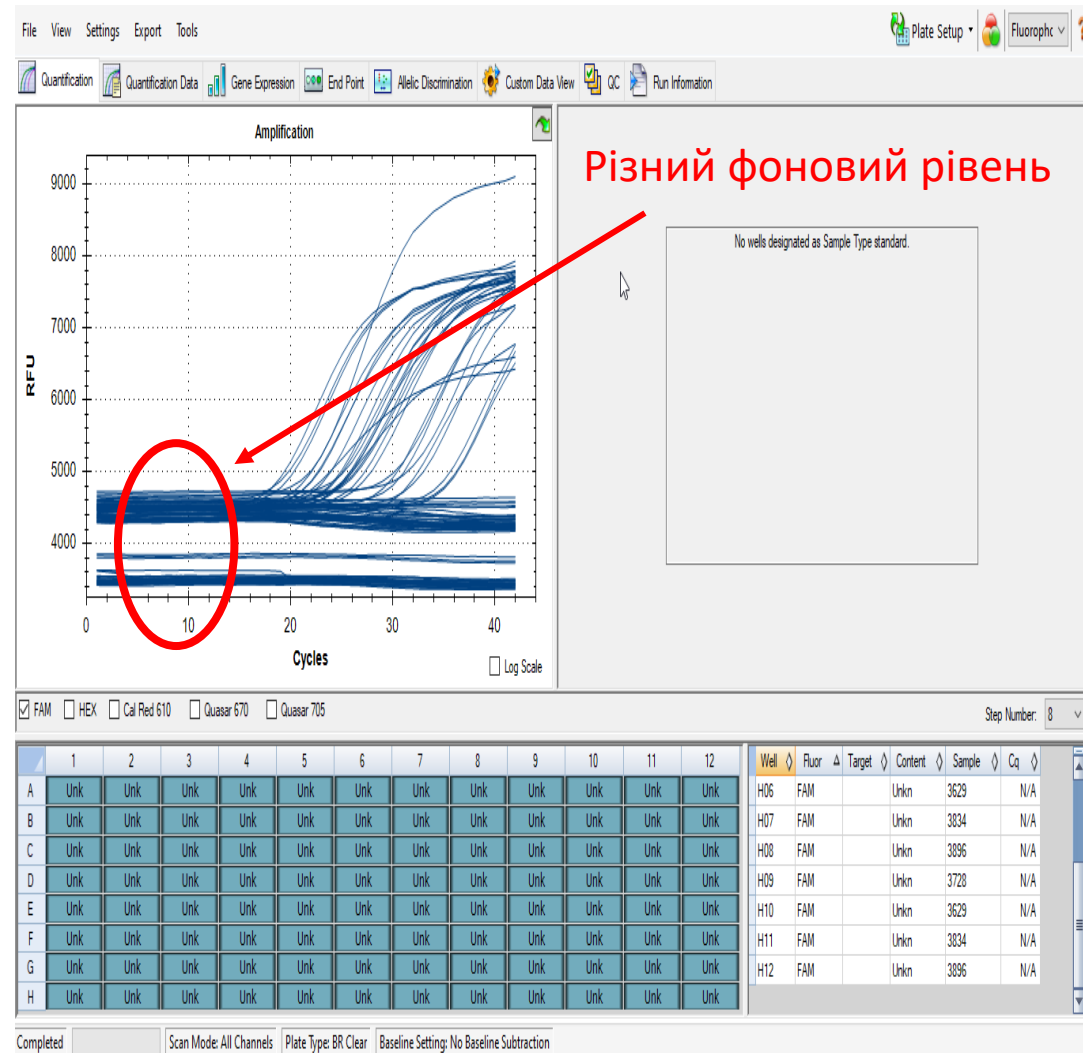
**Вебінар 3.**  
**Пороговий та геометричний методи порівняння графіків ПЛР.**  
**Перша та друга похідна.**  
**Визначення  $S_t$ ,  $S_r$ ,  $S_q$ .**  
**Встановлення порогової лінії**



**Вебінар 4.** Помилки програмного забезпечення ампліфікатора при обробці даних.  
Можливості оператора при роботі з приладами відкритого типу.  
Суцільна практика

# «Сирі» дані ПЛР

# Первинно-оброблені дані



# Первинний процесинг ПЛР-даних

- Масштабування
- Логарифмування
- Віднімання «фону» та зведення ПЛР-кривих на єдину «базову» лінію
- Нормування коефіцієнтів  $\alpha$
- **Лінійна апроксимація\***, згладжування ПЛР-кривих
- Побудова похідних ПЛР-кривих
- Встановлення рівня порогової лінії (для порогового методу аналізу)

**Апроксимація**  
(наближення) – метод, що полягає в заміні одних об'єктів іншими, близькими до вихідних, але більш простими

# Лінійна апроксимація (лінійне фітування)

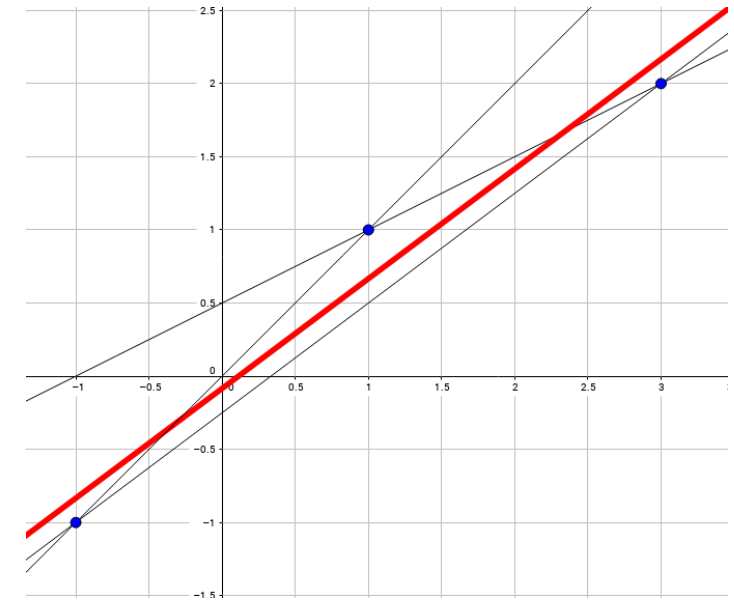
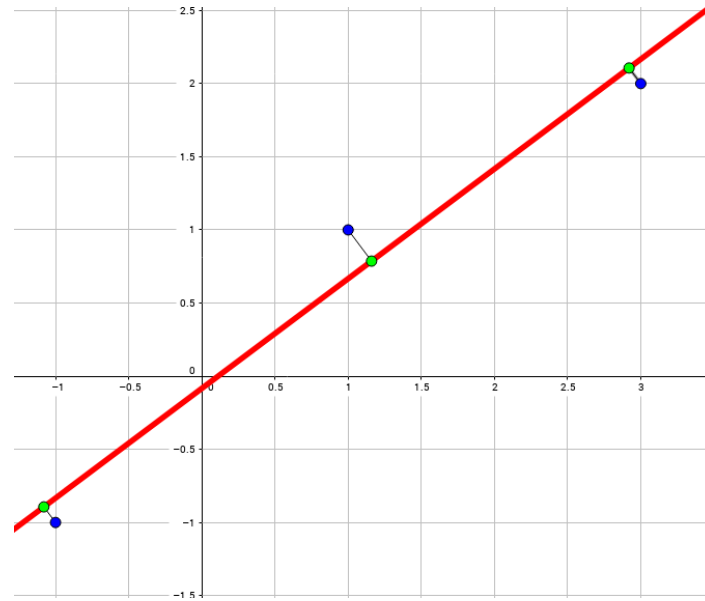
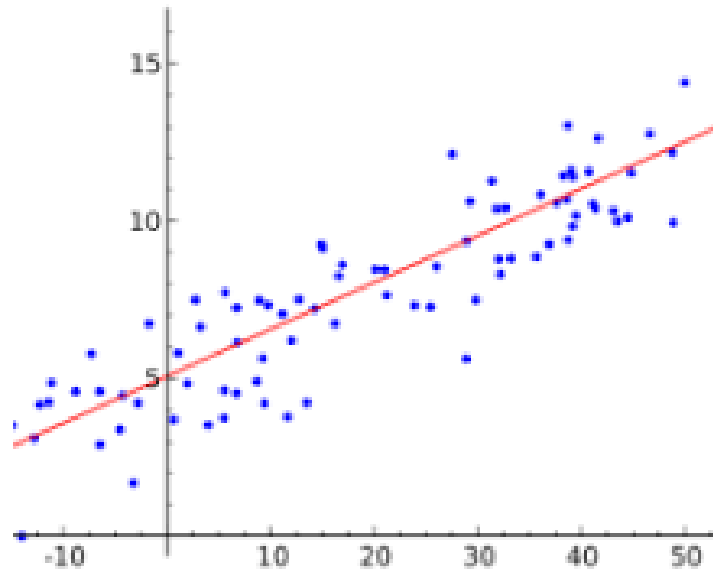
**Фітування** (англ. Fitting) – «припасування», завдяки чому зигзагоподібну фонову фазу ПЛР-кривої перетворюють в пряму лінію, що найбільше підходить для ряду даних.

Досягається за допомогою **лінійної регресії** (лінійного наближення) чи побудови **лінії тренду**.

## Мета:

- Обнуління фонового шуму
- Зменшення інших ділянок ПЛР-кривої на величину фонові флуоресценції, відповідно до лінії тренду фонові флуоресценції.
- Усунення неспецифічних викидів, ігнорування їх при розрахунках, зниження флуоресценції всіх точок графіка ПЛР на амплітуду неспецифічних викидів
- Визначення першого циклу експоненційної фази ПЛР-кривої





## Лінійна регресія та метод найменших квадратів

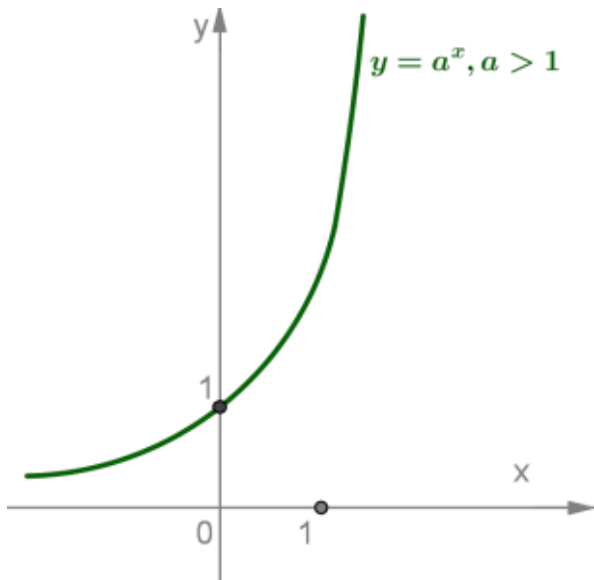
- Лінійна регресія – статистична модель **лінійної** залежності змінної  $y$  від іншої (чи кількох інших) змінної  $x$
- Допомагає припасувати («підігнати») пряму, яка проходить через множину точок
- Один з методів побудови цієї моделі – метод найменших квадратів

# Лінійна регресія та метод найменших квадратів

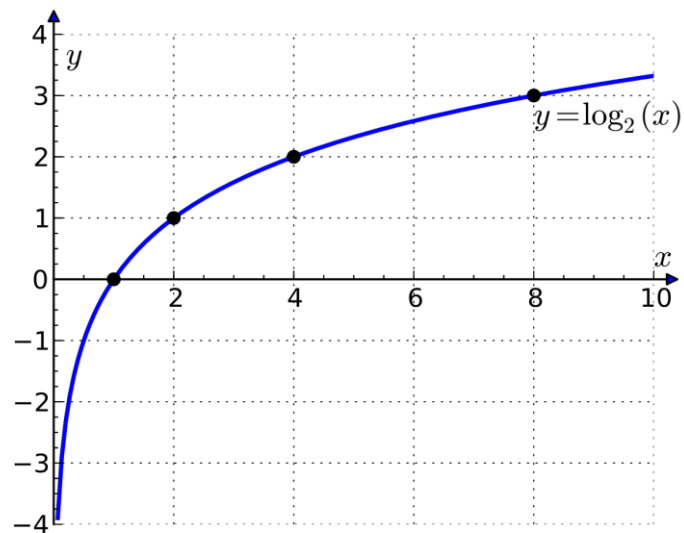
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	Цикл ампліфікації	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
2	Флуоресценція	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11	20	25	30	35	37	40	40	41	40	41	41	41	42	41	43	43	44	45



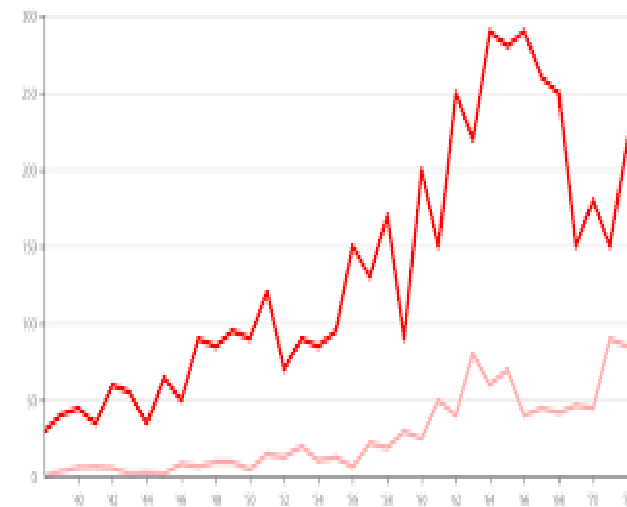
- ПЛР-криві – набір значень послідовного вимірювання рівня флуоресценції в лунці на кожному циклі ампліфікації
- Графік ПЛР описує залежність (тобто, функцію) рівня флуоресценції (залежна змінна **y**) від циклу ампліфікації (змінна **x**)



Експоненційний



Логарифмічний



Лінійний

## Питання 1

Виходячи з усіх наших знань про ПЛР, як ви вважаєте який характер збільшення рівня флуоресценції /кількості ампліфікованих фрагментів?

- Лінійний
- Логарифмічний
- Експоненційний



# Питання 2: Які ділянки графіку ампліфікації більше всього підходять під лінійну модель?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	Цикл ампліфікації	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
2	Флуоресценція	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11	20	25	30	35	37	40	40	41	40	41	41	41	42	41	43	43	44	45



# Дані - позначення

- **Часові ряди** - набір значень, виміряних у часі і розташованих в хронологічному порядку

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	Цикл ампліфікації	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
2	Флуоресценція	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11	20	25	30	35	37	40	40	41	40	41	41	41	42	41	43	43	44	45

- **Змінна** – щось, що може змінювати своє значення. Змінна **y** – **залежна змінна**. Її ми хочемо пояснити, передбачити або зрозуміти, від чого вона залежить. Змінна **x** – **пояснююча змінна** (прогресор).
- По кожній змінній **n** (кількість) спостережень –  **$y_1, y_2, \dots, y_n$**
- **Модель** – це формула, що пояснює змінну і тим самим дозволяє передбачити її (змінної) значення
- **Функція** – співвідношення, відповідність, залежність між змінними

# Лінійна регресія та метод найменших квадратів

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	Цикл ампліфікації	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
2	Флуоресценція	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11	20	25	30	35	37	40	40	41	40	41	41	41	42	41	43	43	44	45



- Формула лінійної моделі

$$y = b1 + b2 * x + e,$$

Де  $x$ ,  $y$  – змінні ( $y$  – рівень флуоресценції,  $x$  – цикл ампліфікації)

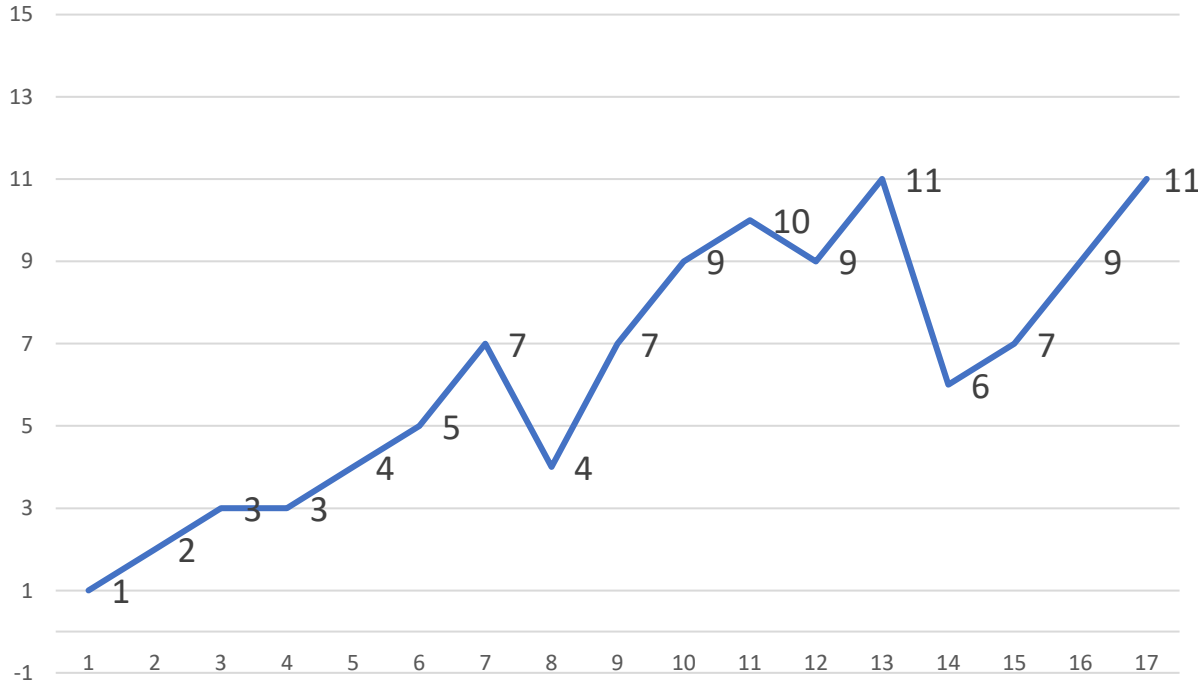
$b1$ ,  $b2$  – невідомі коефіцієнти.  $b2$  показує на скільки зміниться рівень флуоресценції на наступному циклі ампліфікації

$e$  – випадкова складова, яку неможливо передбачити, але яка впливає на рівень флуоресценції (напр., помилка піпетування).

# Лінійна регресія та метод найменших квадратів

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Цикл ампліфікації (x)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	Флуоресценція (y)	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11

Флуоресценція (y)



Цикл ампліфікації (x)

- Маємо 17 точок даних з координатами (x;y) – (1;1), (2;2), (3;3), (4;3), (5;4), (6;5), (7;7) тощо.

- Підставляємо ці дані в формулу  $y = b1 + b2*x$

$$1 = b1 + b2*1$$

$$2 = b1 + b2*2$$

$$3 = b1 + b2*3$$

$$4 = b1 + b2*5$$

$$5 = b1 + b2*6$$

$$6 = b1 + b2*7 \dots$$

- Розраховуємо похибку прогнозування для кожного рівняння (різницю між правою та лівою частиною рівнянь):

$$1 - (b1 + b2*1)$$

$$2 - (b1 + b2*2)$$

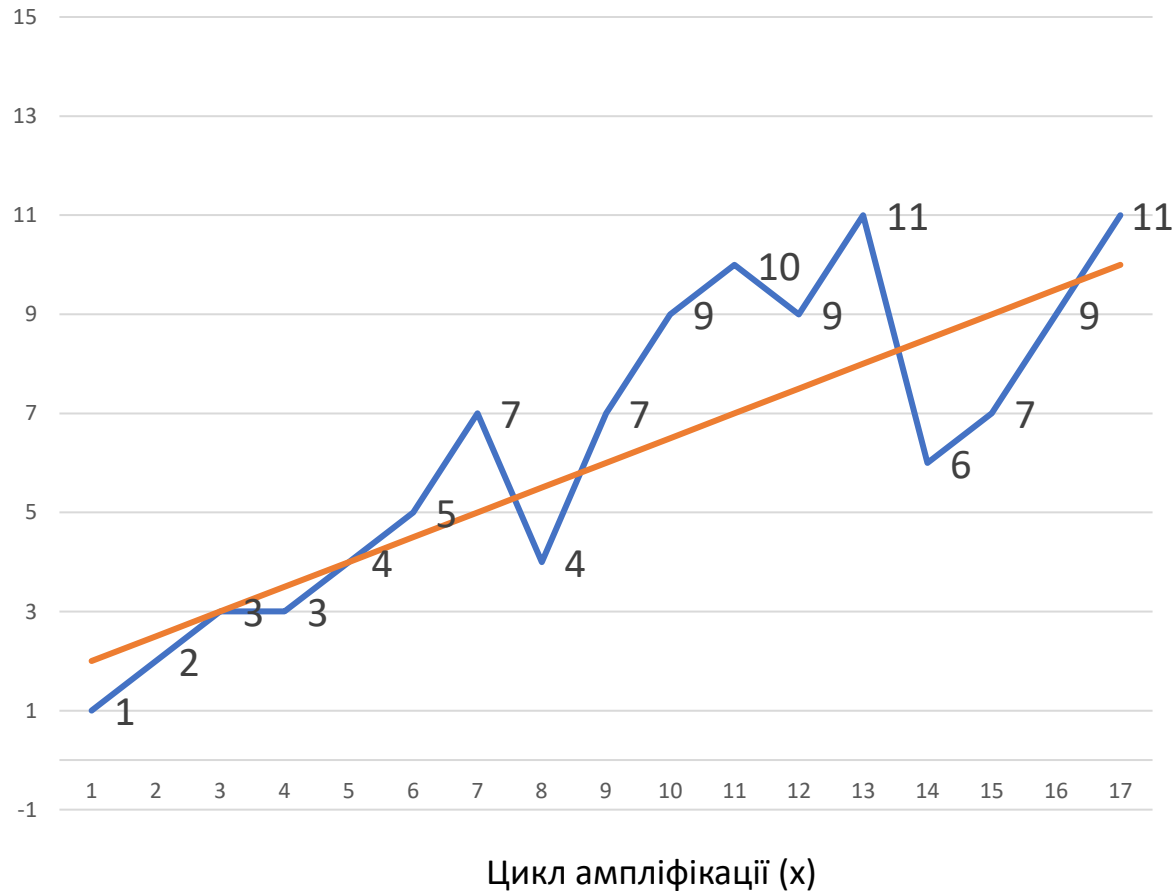
$$3 - (b1 + b2*3) \dots$$

- Підводимо похибки прогнозування в квадрат та розраховуємо їх суму (**S**):

$$S_{(b1,b2)} = [1 - (b1 + b2*1)]^2 + [2 - (b1 + b2*2)]^2 + [3 - (b1 + b2*3)]^2 + \dots$$

# Лінійна регресія та метод найменших квадратів

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Цикл ампліфікації (x)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	Флуоресценція (y)	1	2	3	3	4	5	7	4	7	9	10	9	11	6	7	9	11
4	Ампроксимовані дані (y)	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10



- Знаходимо мінімум функції двох змінних  $b_1$  і  $b_2$ .

$$\frac{\partial S(b_1, b_2)}{\partial b_1} = 0$$

$b_1$

та

$$\frac{\partial S(b_1, b_2)}{\partial b_2} = 0$$

$b_2$

Отримуємо рівняння з двома невідомими, розв'язком для якого буде:

$$b_1 = 1.5$$

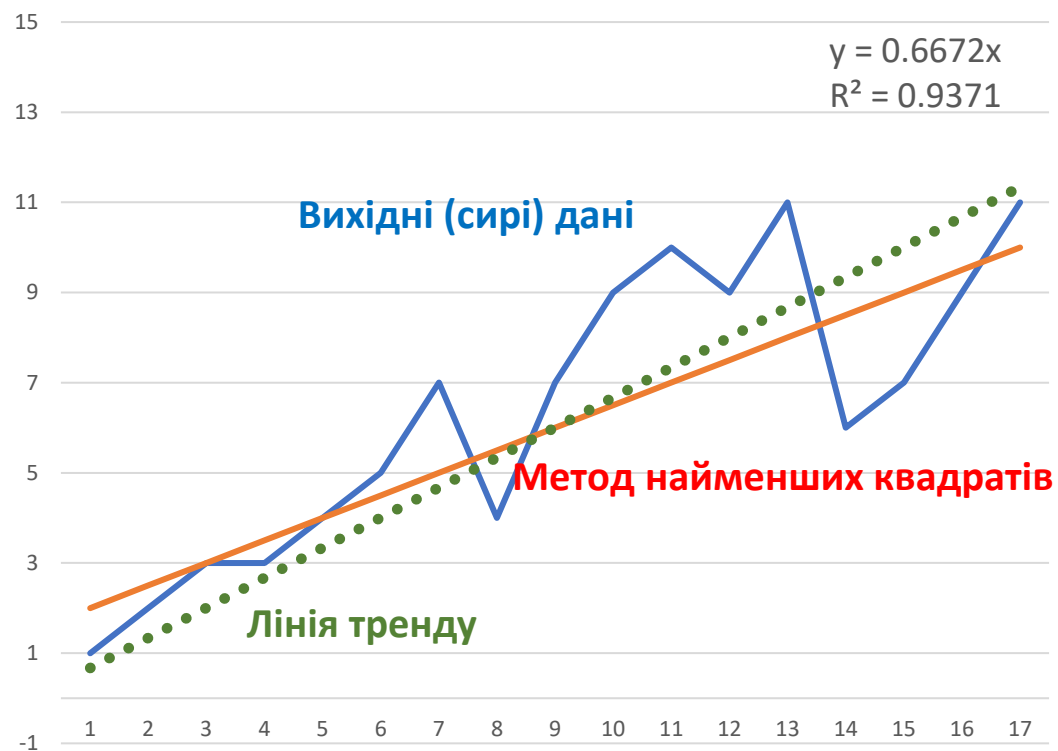
$$b_2 = 0.5$$

Підставляємо коефіцієнти в рівняння:

$$y = 1.5 + 0.5 * x,$$

що є рівнянням прямої, яка проходить найближче до поданих 17 точок

# Як це зробити в Excel

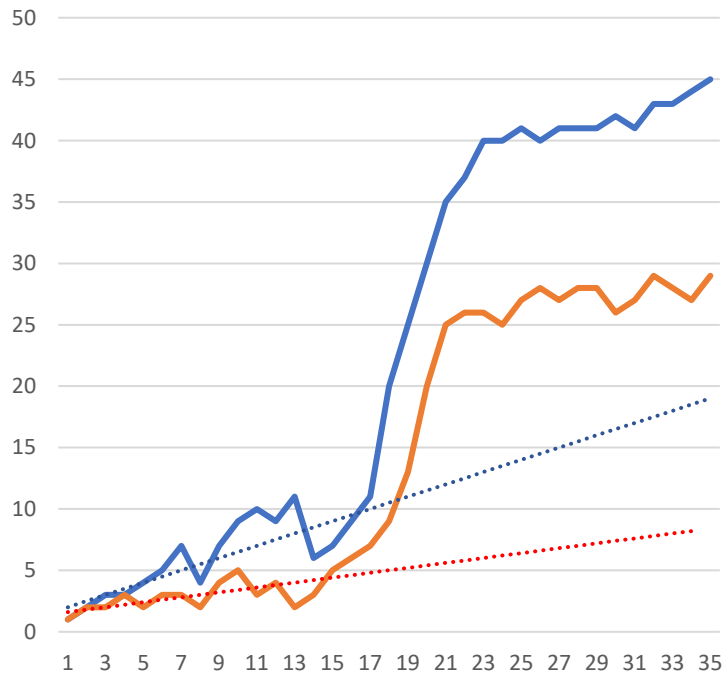


1. Використати функцію =LINEST() / =ЛИНЕЙН() для масиву даних,
2. Побудувати лінію тренду

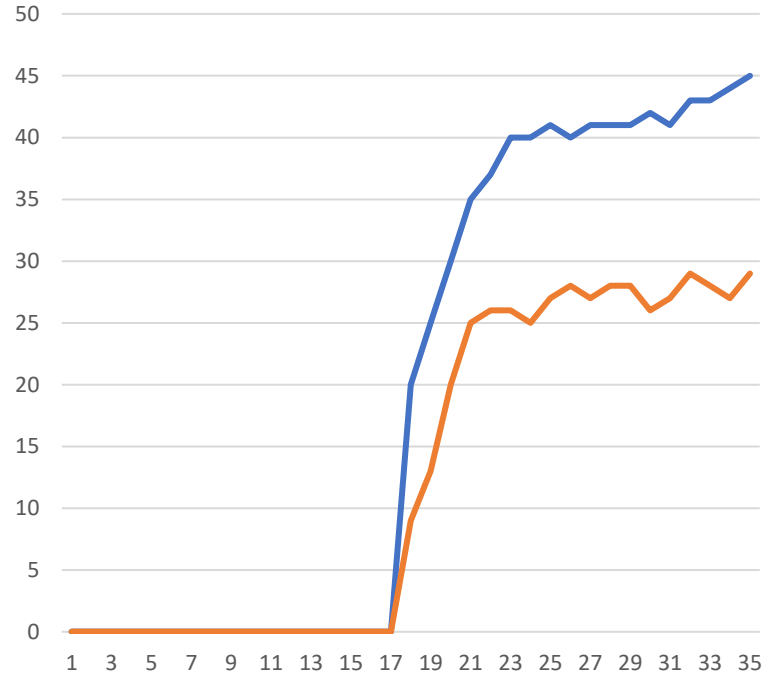
**Лінія тренду** – це візуальне відображення напрямку змін значень якогось показника в часі (характер, тенденція та динаміка змін).

Уявляє собою геометричне відображення середніх значень показників, що аналізуються

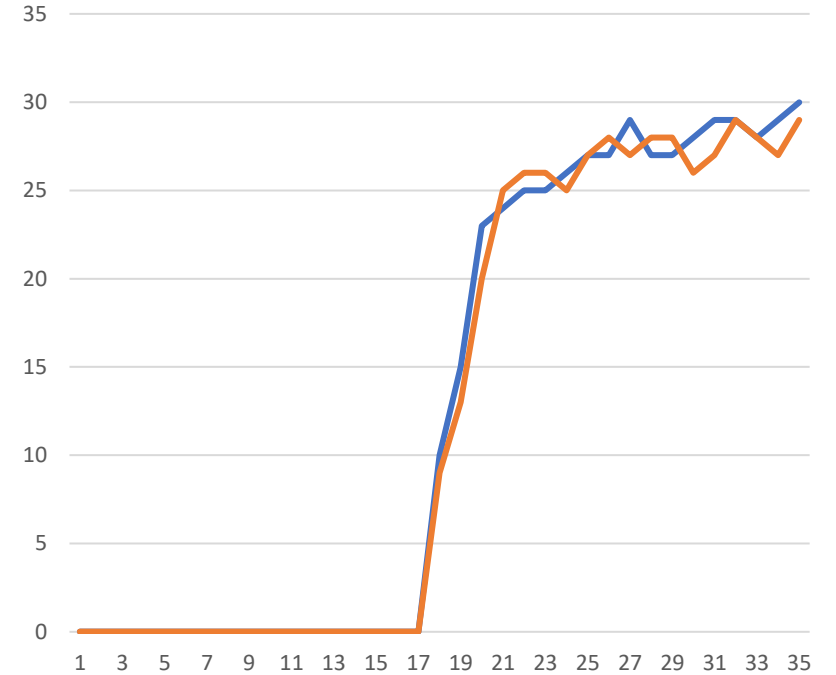
# Мета №1 апроксимації – обнуління фонового шуму



Графік 1

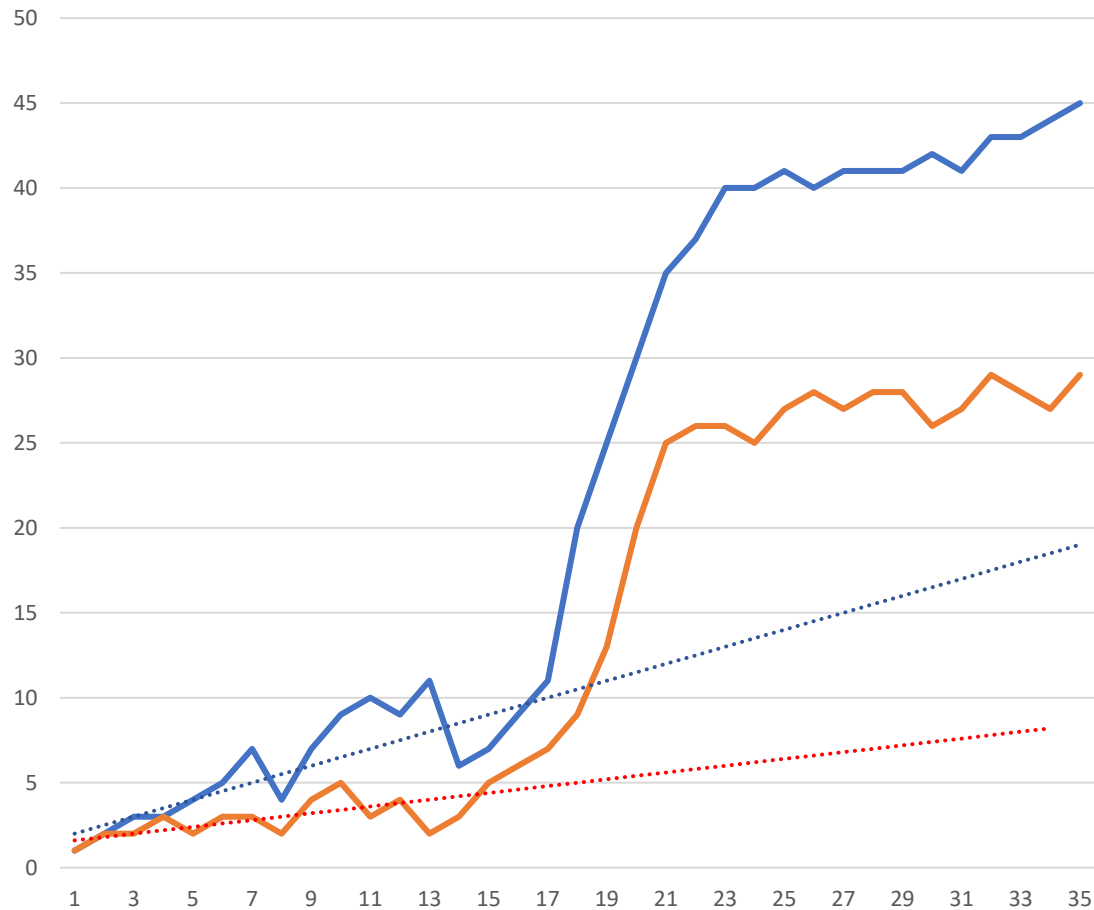


Графік 2

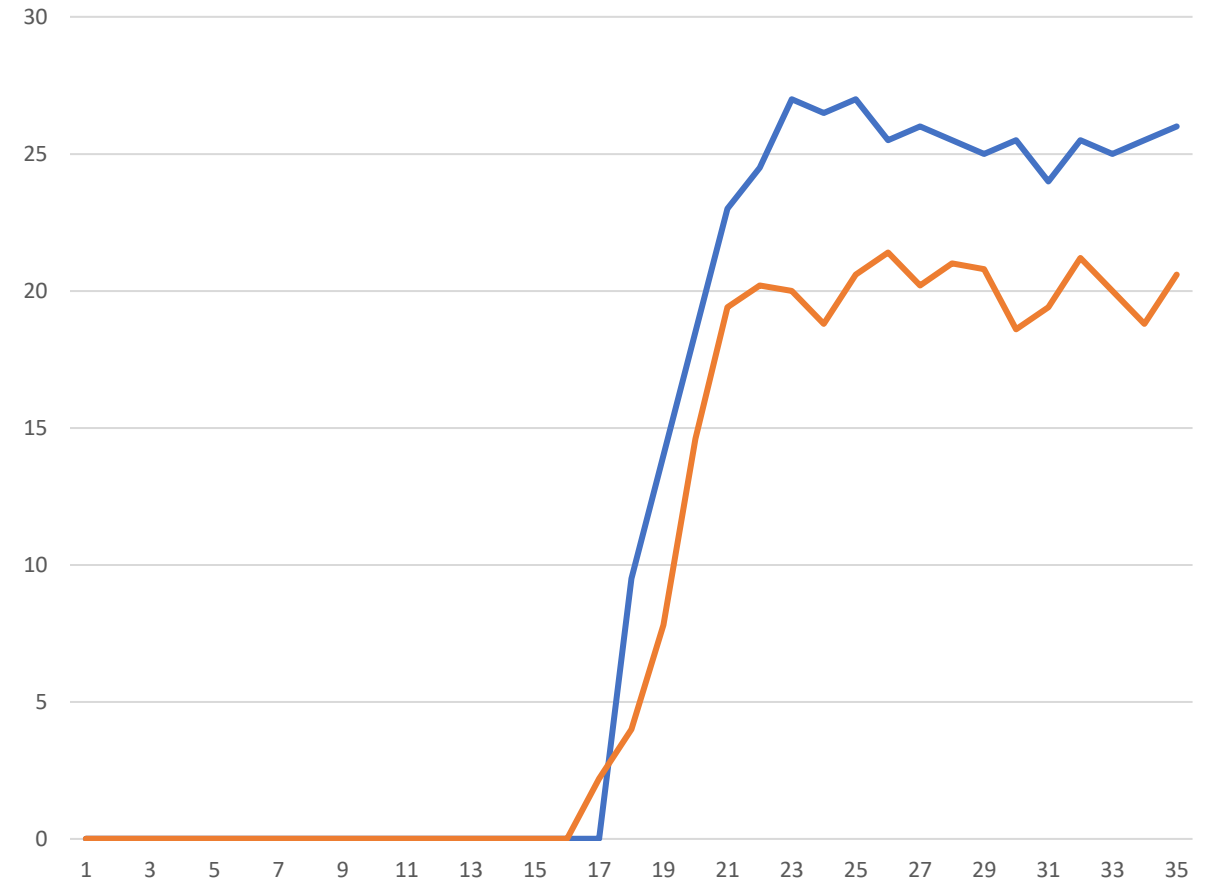


Графік 3

# Мета №2 апроксимації – зменшення інших ділянок кривої на фонову флуоресценцію



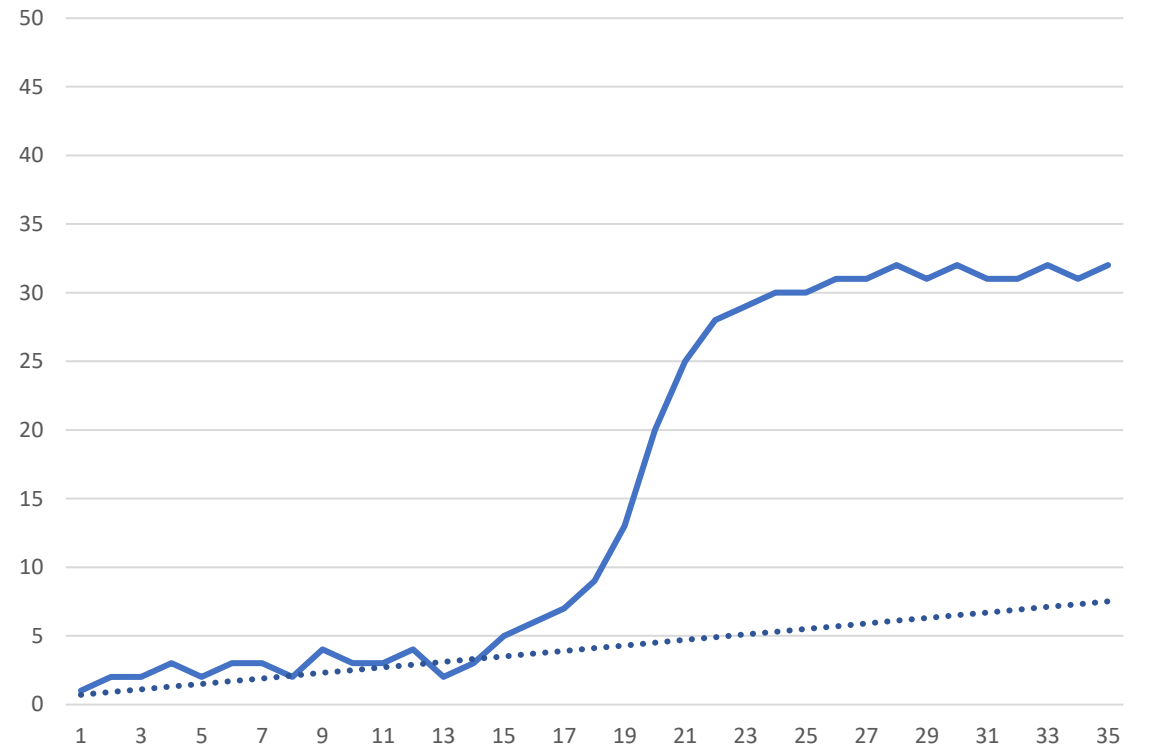
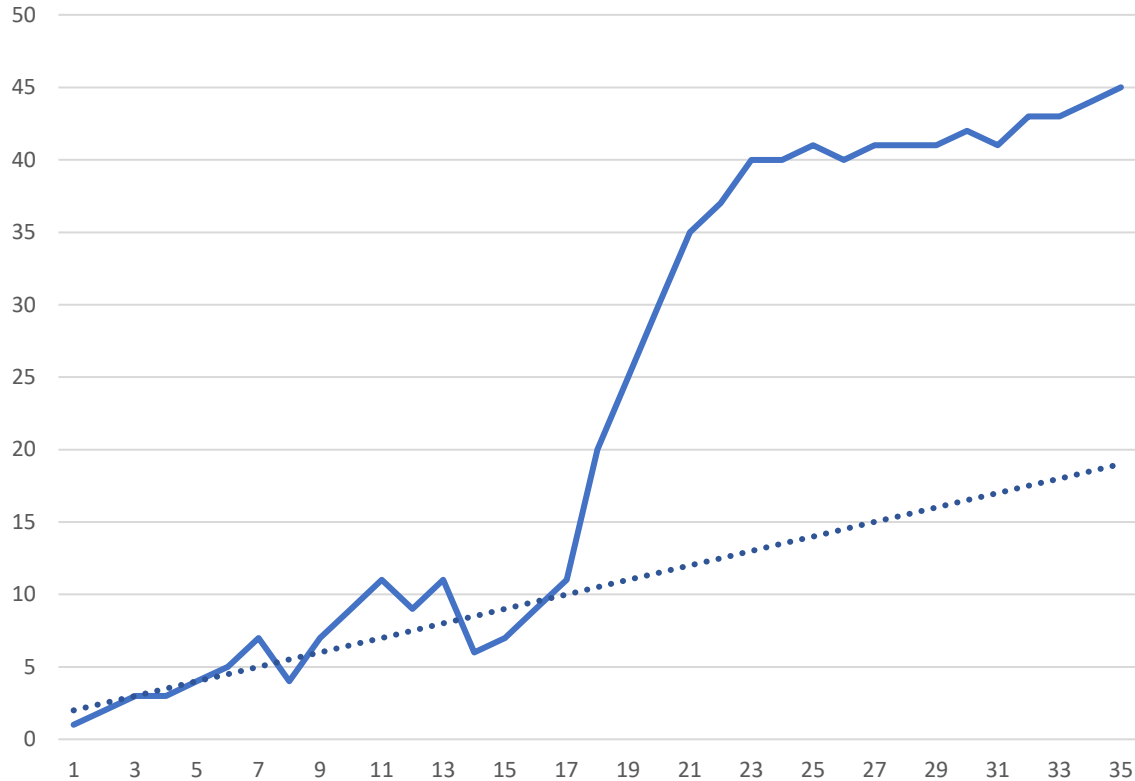
До віднімання фону



Після віднімання фону

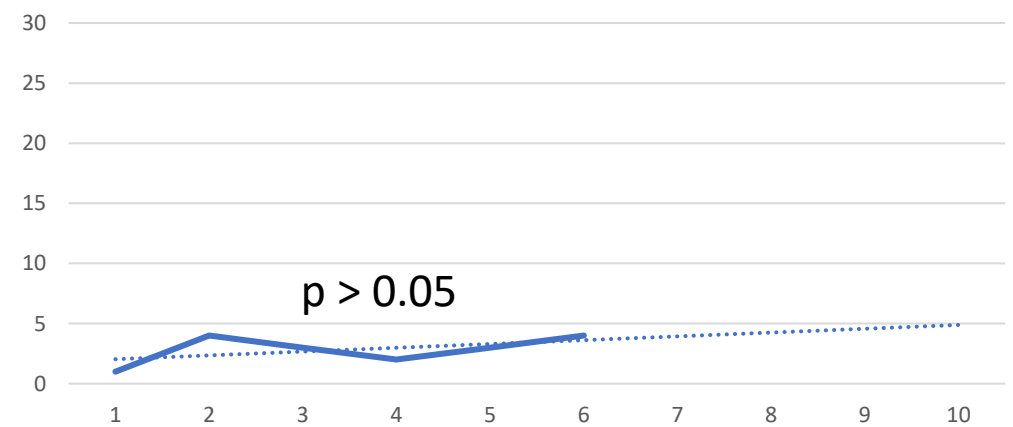
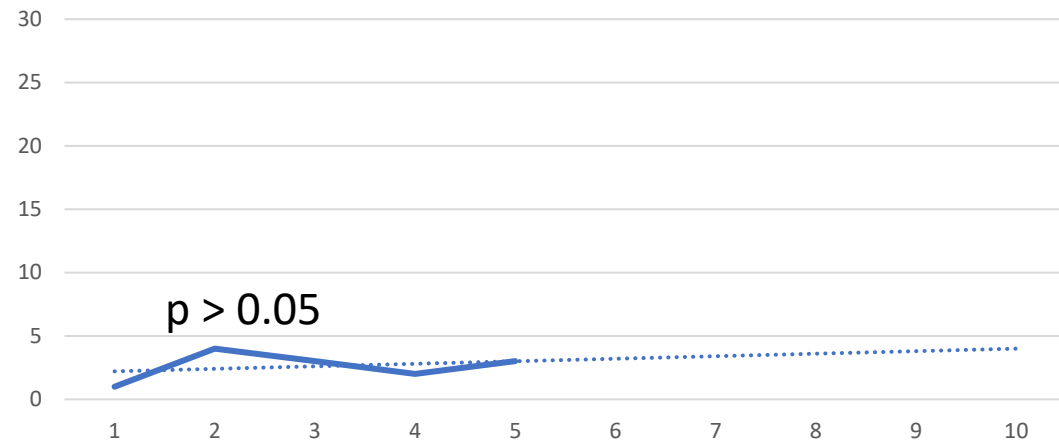
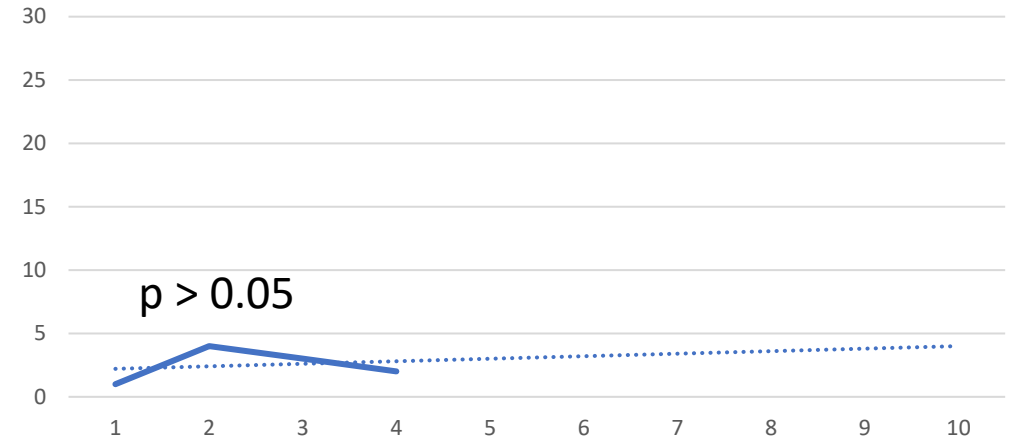
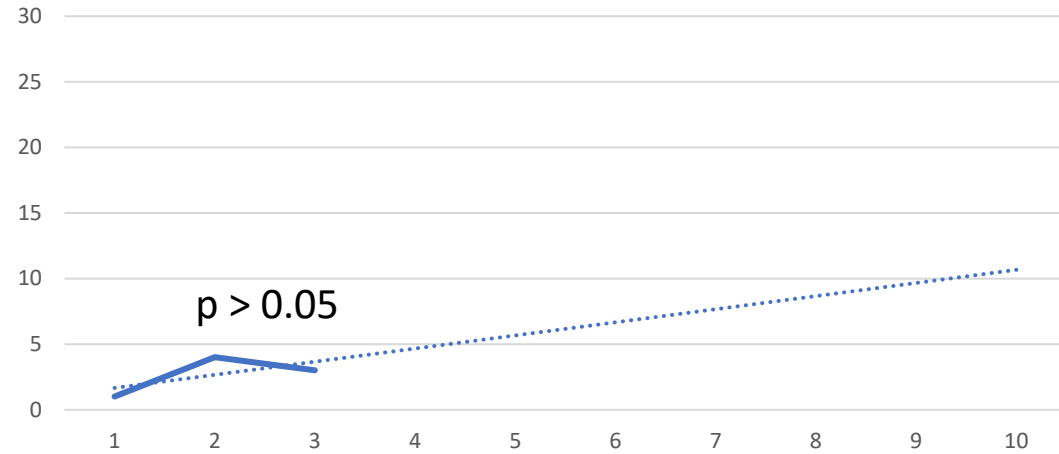


# Мета №3 апроксимації – виключення неспецифічних викидів

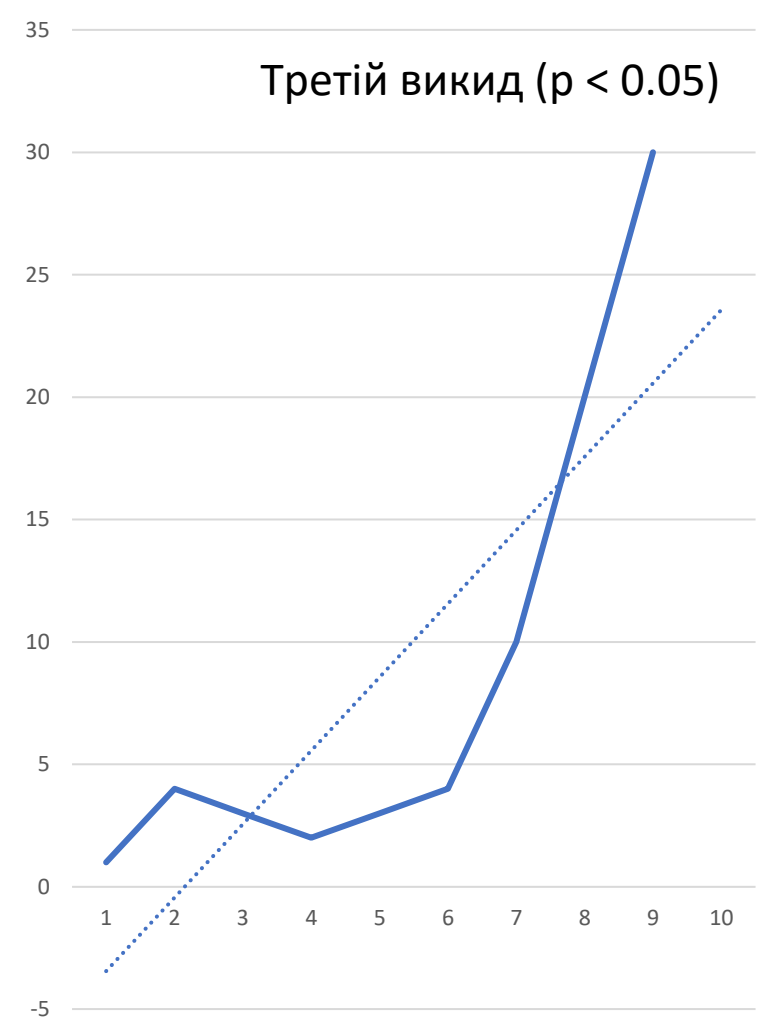


Неспецифічний викид усунений

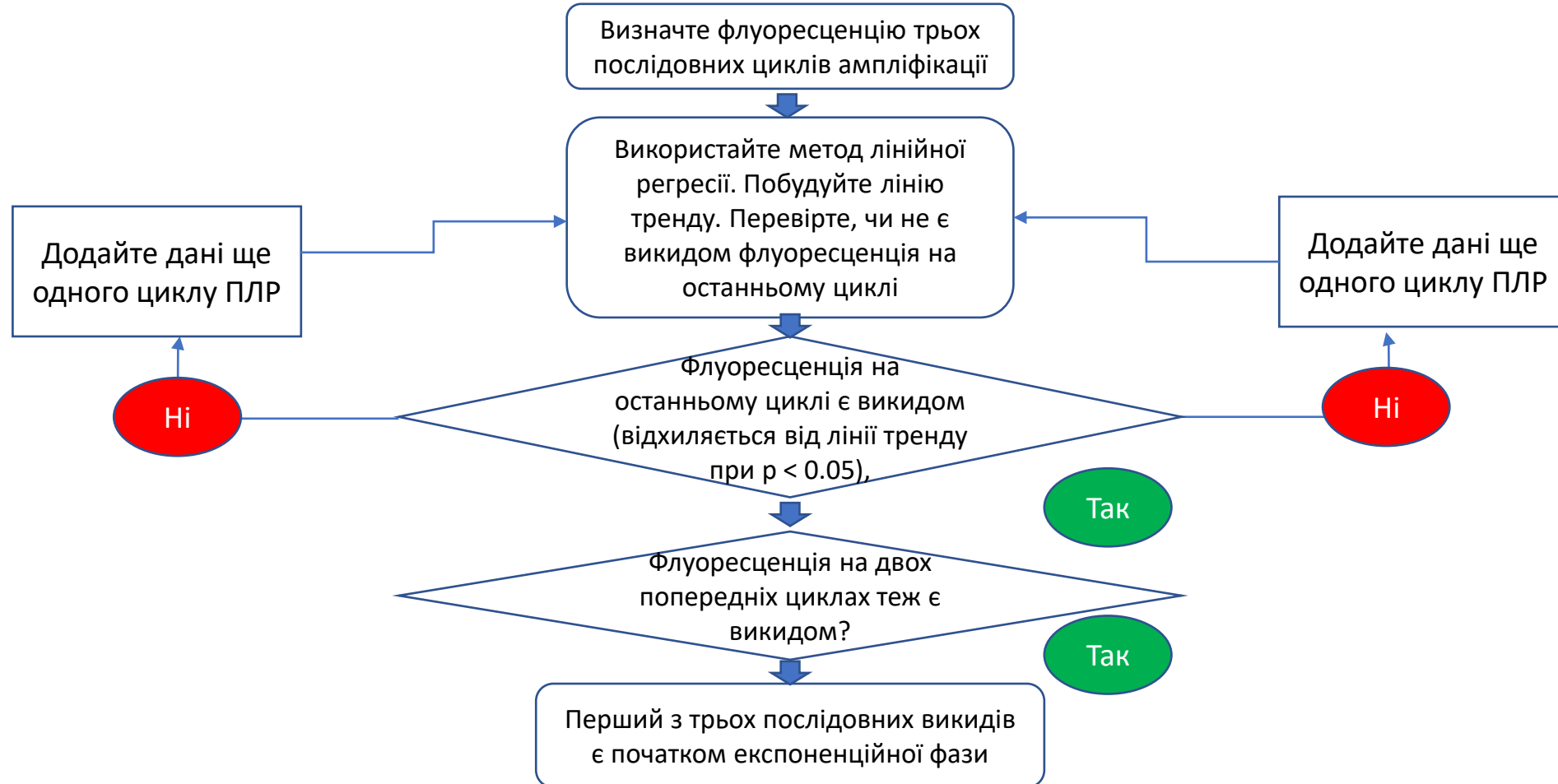
# Мета №4 апроксимації – визначення першого циклу експоненційної фази



# Мета №4 апроксимації – визначення першого циклу експоненційної фази



# Мета №4 апроксимації – визначення першого циклу експоненційної фази



Дякую за увагу

---