



Animal Health Matters.  
For Safe Food Solutions.



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,  
Education and Research EAER

**State Secretariat for Economic Affairs SECO**



# Міри поширення захворювання

Марко Де Нарді

# Цілі

1. Описати такі міри поширення хвороби: інтервал, міжквартильний діапазон, дисперсія і стандартне відхилення
2. Обговорити приклади використання мір поширення захворювання в епідеміології

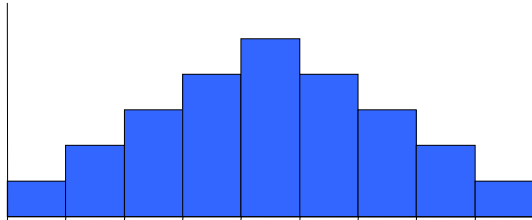


# Ціль

Описати такі міри поширення хвороби: інтервал, міжквартильний діапазон, дисперсія і стандартне відхилення

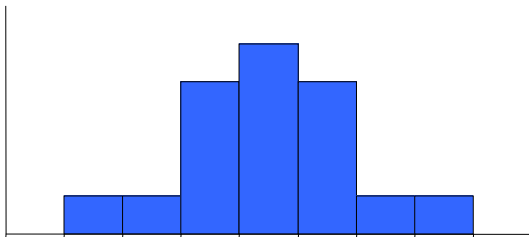


# Розуміння поняття “поширення”

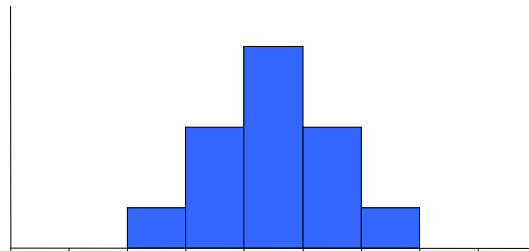


Однаковий центр

Вік у 3  
групах  
студентів



але ...

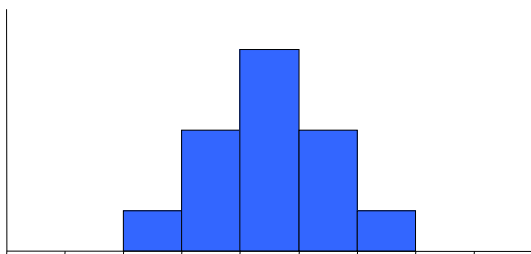
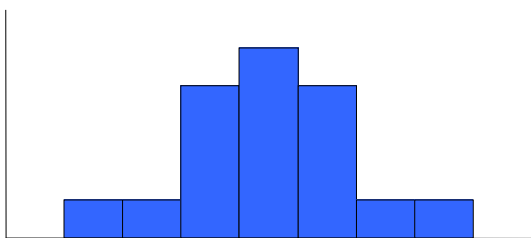
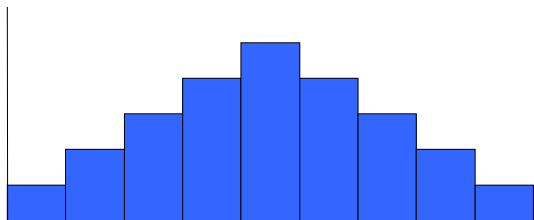


різне поширення



# Розуміння поняття “поширення”

Вік у 3  
групах  
студентів



**Частотний розподіл** – це представлення даних або у графічному або у табличному форматі, і демонструє спосіб розподілення даних по діапазону значень, показуючи число і відсоток осіб в межах групи значень.

В цьому випадку всі вони мають високі частоти в центрі розподілу і низькі частоти на 2 екстремумах (**верхній і нижній** краї розподілу)



# Міри поширення захворювання

**Визначення: кількісно визначають суму варіацій, або дисперсії в наборі даних**

Також відомі як:

- “міра дисперсії”
- “міри варіації”

Загальні міри

- діапазон
- міжквартильний діапазон
- дисперсія
- стандартне відхилення



# Діапазон

Визначення: різниця між найбільшими і найменшими значеннями

- **У статистиці** діапазон подається як одне число і є результатом віднімання максимального значення від мінімального
- **В епідеміології**, діапазон – це найбільші і найменші спостереження в зразку, а не різниця між найбільшим і найменшим значенням



# Визначення діапазону тривалості перебування

Кількість ночей, які провели 27 пацієнтів у лікарні після зараження *Staphylococcus spp.*...

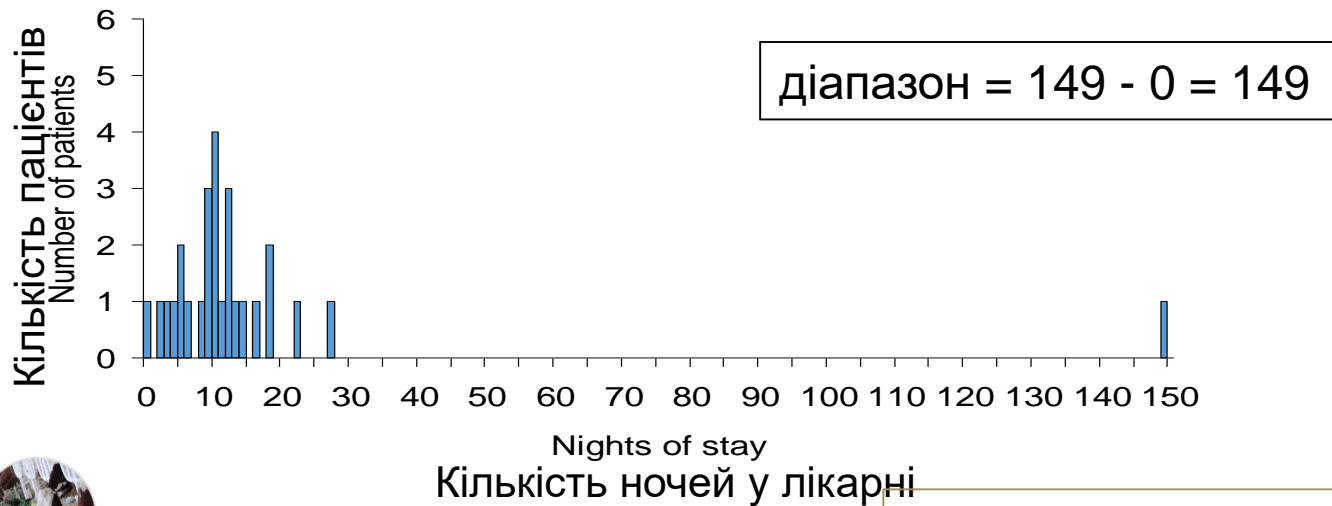
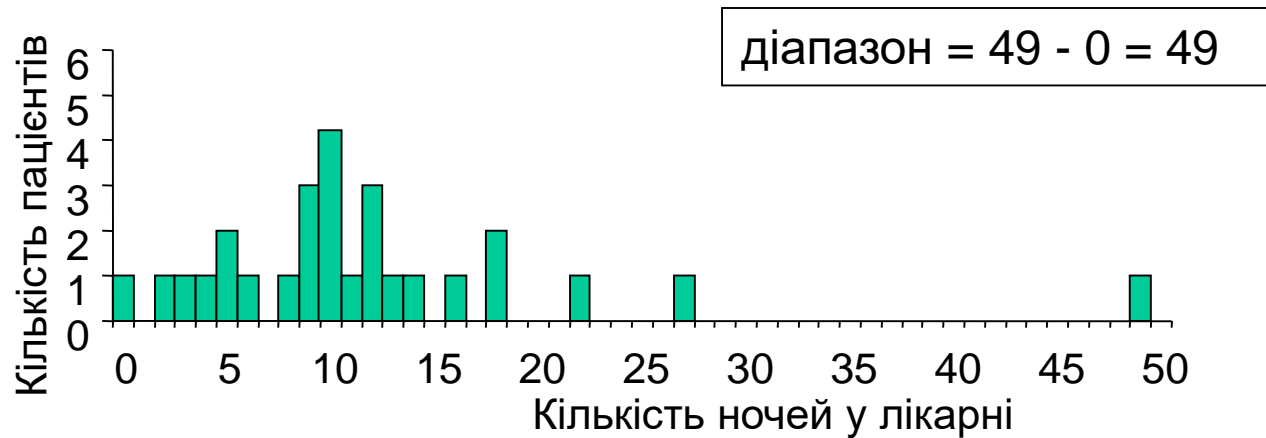
0, 2, 3, 4, 5, 5, 6, 8, 9,  
9, 9, 10, 10, 10, 10, 11, 12, 12,  
12, 13, 14, 16, 18, 18, 22, 27, 49

Який діапазон значень?





# Чи чутливий діапазон до крайніх значень (екстремальних значень, випадів)?



# Міжквартильний діапазон (IQR)

Визначення: **Різниця між першим і третім "квартилем"** (визначення на наступному слайді)

Властивості / Область застосування:

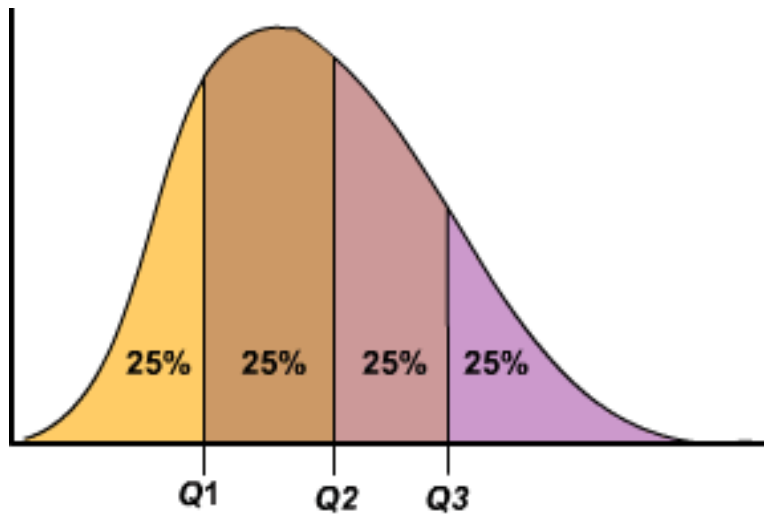
- Модифікація діапазону, який менш чутливий до випадів
- Розраховується як різниця між двома значеннями даних, але не двох крайніх значень



# Квартилі

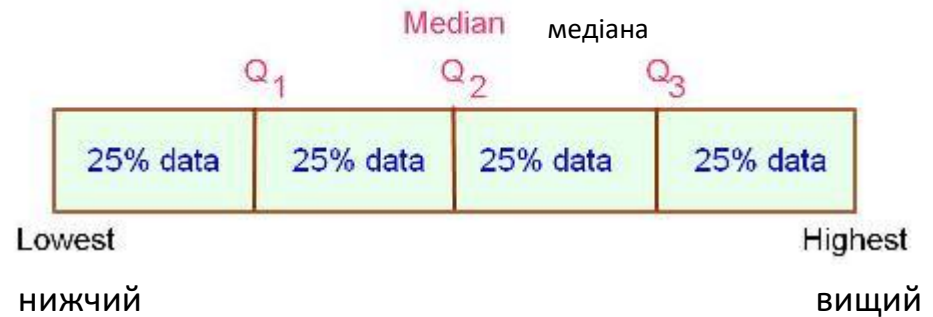
- **Визначення: Квартилі – це значення, які поділяють дані на чотири рівні частини**
- 25% спостережень нижче першого квартиля (Q1)
- 25% спостережень знаходяться між Q1 і Q2 (медіана)
- 25% спостережень знаходяться між Q2 (медіана) і Q3
- 25% спостережень знаходяться вище Q3



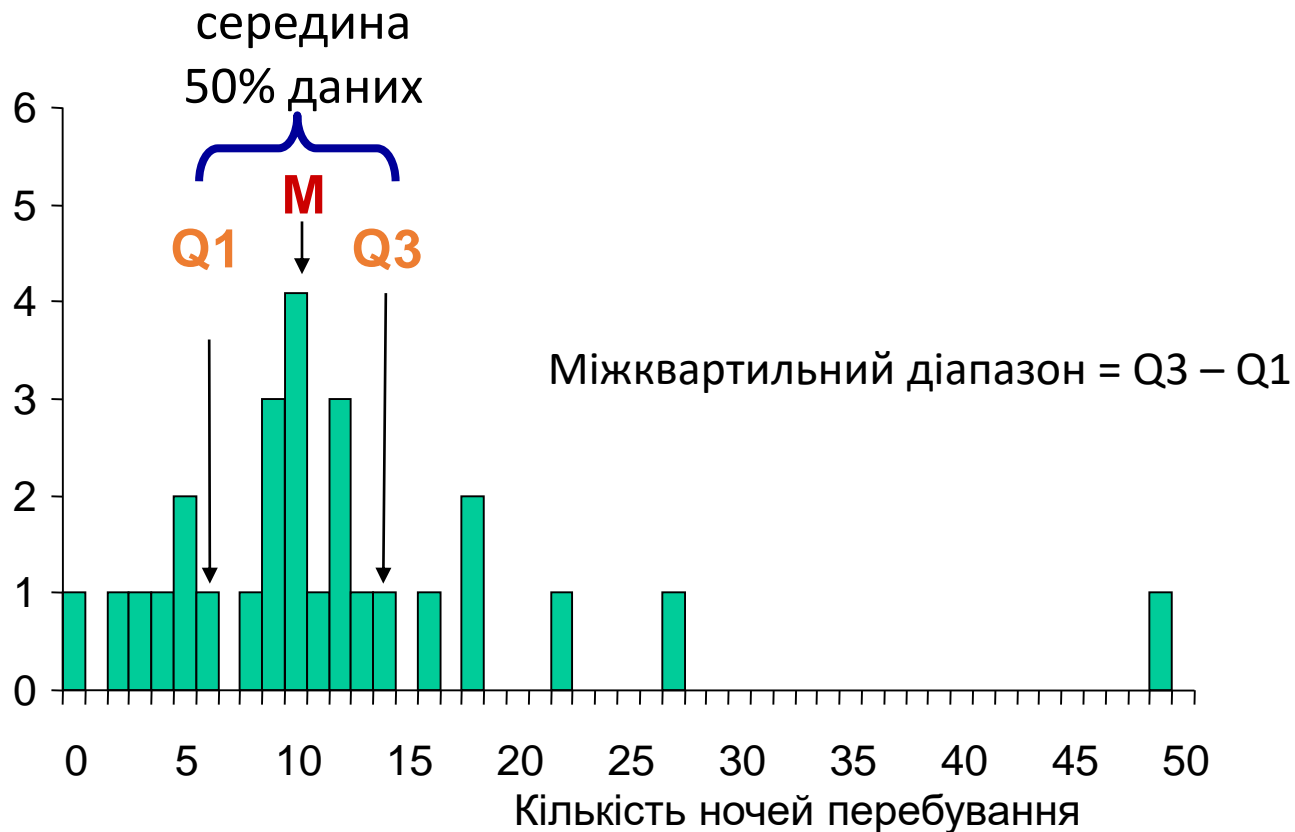


Квартилі і розподіл даних

Quartiles and data distribution



# Міжквартильний діапазон — дані по тривалості перебування



# Розрахунок позицій кватилів у наборі даних з $N$ спостережень

Позиція Q1:	$1 * (n+1) / 4$
Позиція (медіана):	$2 * (n+1) / 4$
Позиція Q3:	$3 * (n+1) / 4$

-----  
Приклад: У базі даних з 55 спостереженнями, позиція третього кватилля (Q3) буде:

$$3 * (55+1) / 4 = 168 / 4 = 42$$

Таким чином, третій кватиль є значення в 42-й позиції в упорядкованому списку



# Визначення міжквартильного діапазону даних по тривалості перебування

Кількість ночей, які провели 27 пацієнтів у лікарні

0, 2, 3, 4, 5, 5, **Q1** 6, 8, 9,  
9, 9, 10, **Q2** 10, 10, 10, 11, 12, 12,  
12, 13, 14, 16, 18, 18, 22, 27, 49

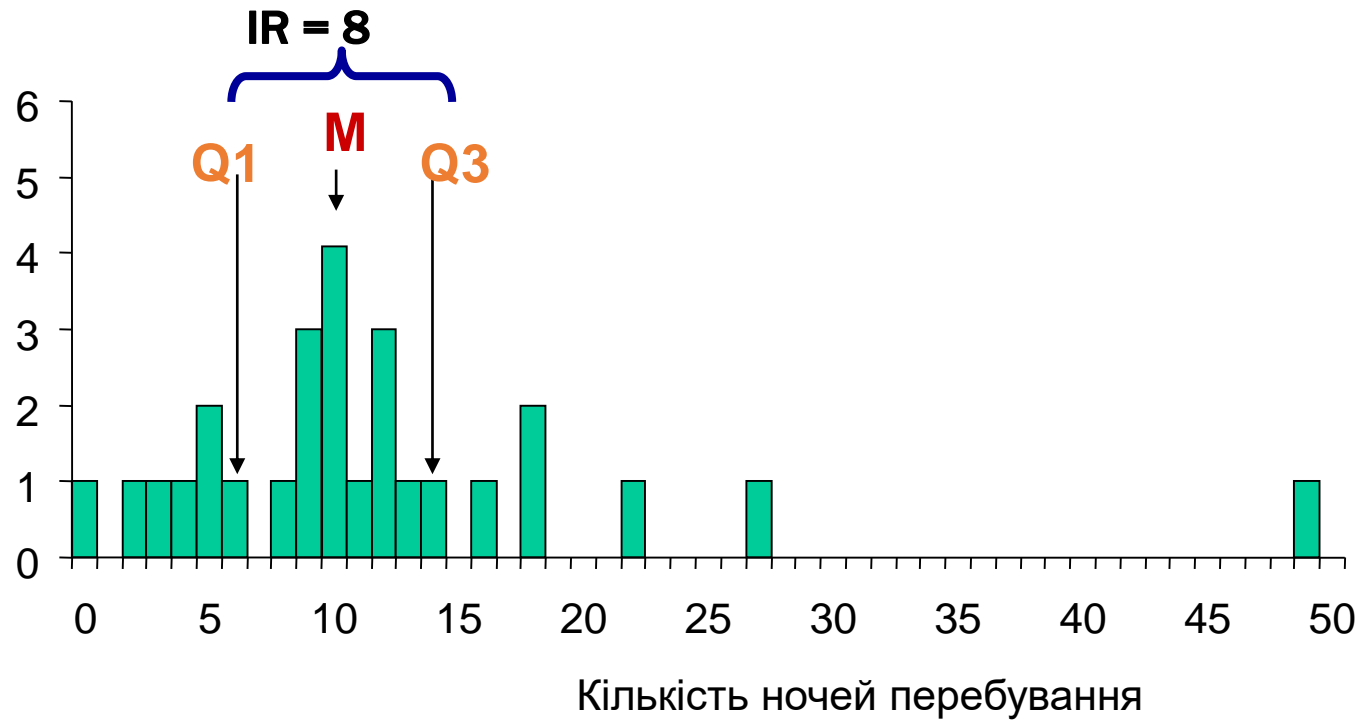
**Q3**

	<u>Позиція</u>	<u>Значення</u>
<b>Q1:</b>	$1 * (27 + 1) / 4 = 7^{\text{th}}$	→ <b>6</b>
<b>Q2 (Median):</b>	$2 * (27 + 1) / 4 = 14^{\text{th}}$	→ <b>10</b>
<b>Q3:</b>	$3 * (27 + 1) / 4 = 21^{\text{st}}$	→ <b>14</b>

Міжквартильний діапазон:  $14 - 6 = 8$



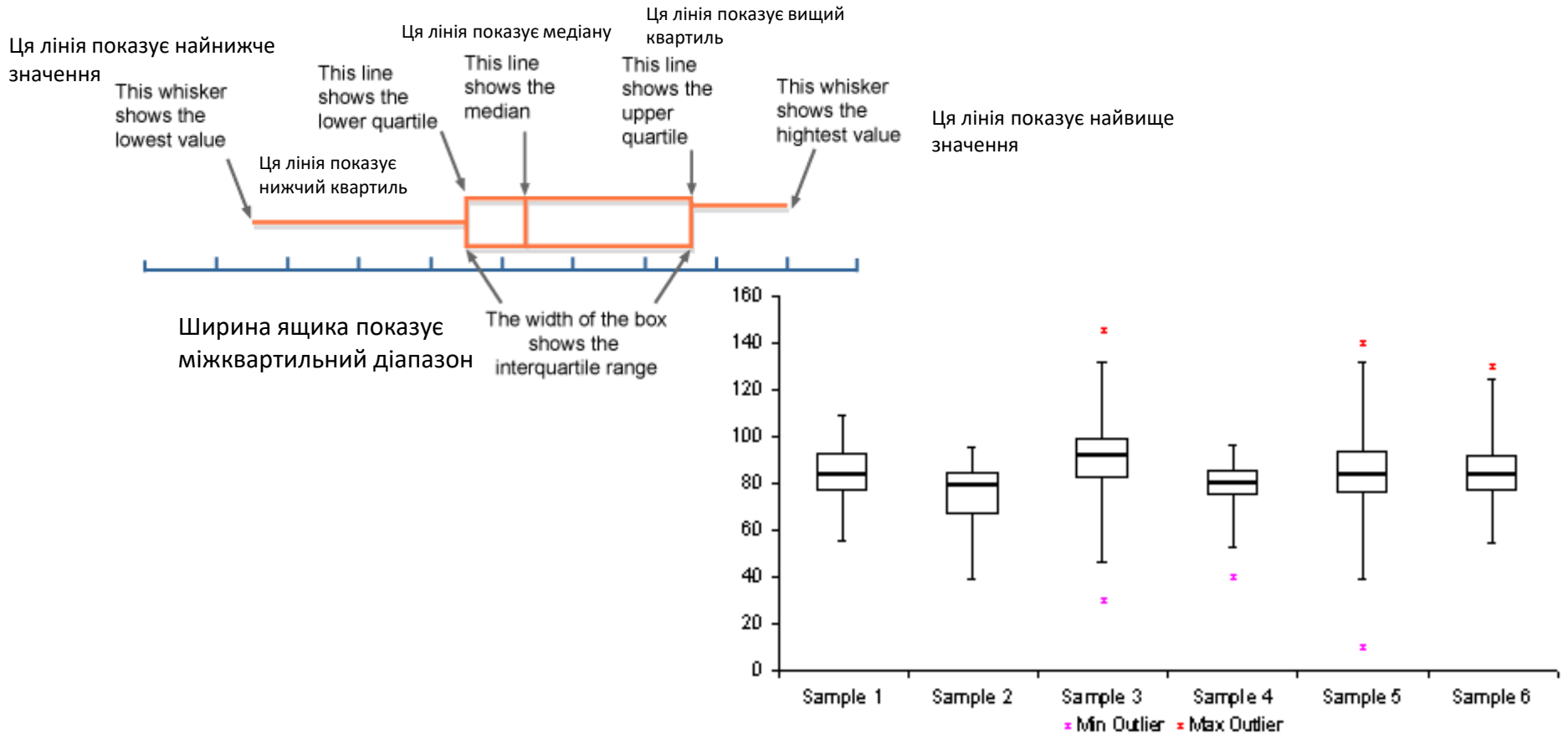
# Міжквартильний діапазон — дата тривалості перебування





# Як представити квартилі й міжквартилі?

## діаграма виду "ящик з вусами"



Мінімальний і максимальний випад



# Дисперсія і стандартне відхилення

Загальний опис: відхилення кожного спостережуваного зразка від середнього значення.

## Дисперсія

- різниця середнього значення квадрату девіації від середнього значення
- сума  $(x_i - \text{середнє значення})^2 / (n-1)$

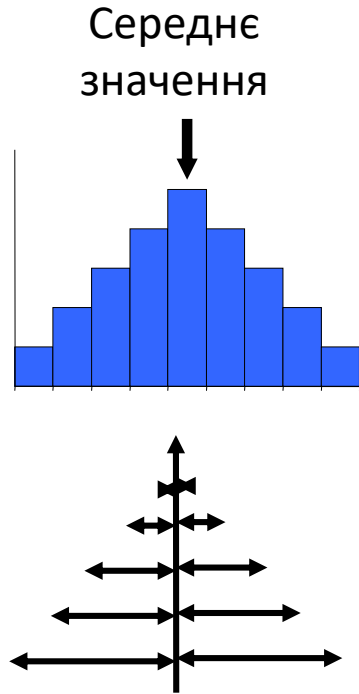
## Стандартне відхилення

- Квадратний корінь дисперсії

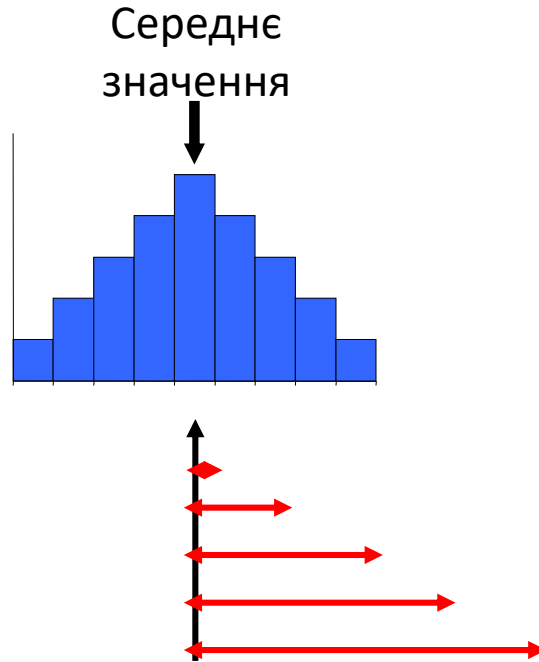
Дисперсія і стандартне відхилення – пов'язані міри, які кількісно обраховують наскільки тісно згруповані спостережувані значення до середнього значення.



# Дисперсія і стандартне відхилення



При описі наскільки кожне спостереження відалене від середнього значення, деякі значення будуть позитивними, а деякі з них будуть негативними



**Дисперсія і стандартне відхилення**  
описати цю дисперсію даних шляхом зведення в квадрат значення для уникнення негативних чисел



# Рівняння для дисперсії і стандартного відхилення

- $\bar{x}$  : середнє значення
- $x_i$  : значення
- $n$  : кількість значень
- $s^2$ : дисперсія
- $s$  : стандартне відхилення

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{(n-1)}$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}}$$



# Стандартне відхилення– Властивості/Область застосування

Коли дані приблизно нормально розподілені (дзвіноподібна крива), то:

- 68,3% даних потрапляють в 1 SD (стандартне відхилення) від середнього значення
- 95.5% даних потрапляють в 2 SD від середнього значення
- 95.0% даних потрапляють в 1.96 SD від середнього значення
- 99.7% даних потрапляють в 3 SD від середнього значення



# Дані тривалості перебування

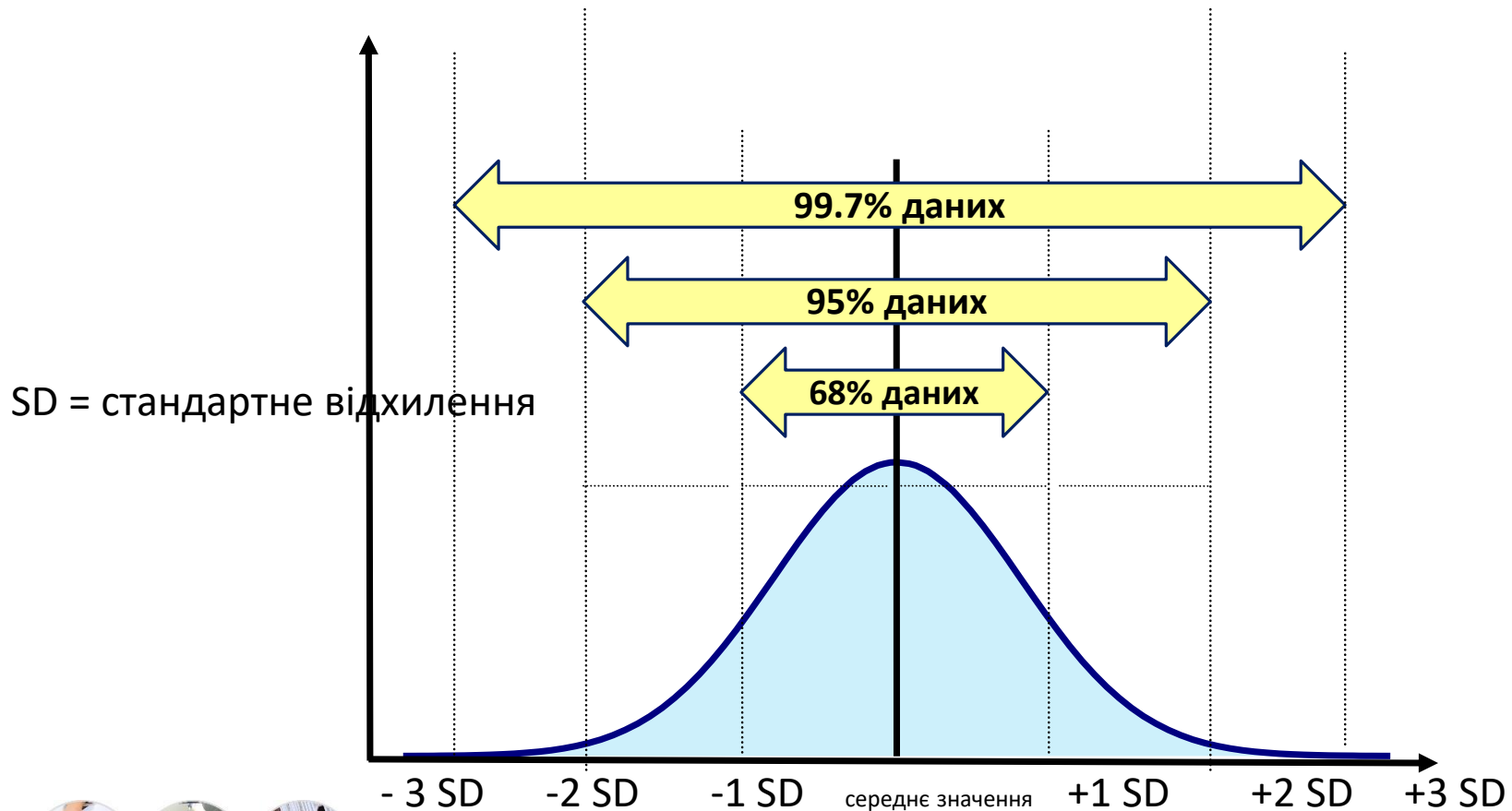
Кількість ночей, які 27 пацієнтів провели у лікарні

$(0 - 12)^2 = 144$	$(9 - 12)^2 = 9$	$(12 - 12)^2 = 0$
$(2 - 12)^2 = 100$	$(9 - 12)^2 = 9$	$(13 - 12)^2 = 1$
$(3 - 12)^2 = 81$	$(10 - 12)^2 = 4$	$(14 - 12)^2 = 4$
$(4 - 12)^2 = 64$	$(10 - 12)^2 = 4$	$(16 - 12)^2 = 16$
$(5 - 12)^2 = 49$	$(10 - 12)^2 = 4$	$(18 - 12)^2 = 36$
$(5 - 12)^2 = 49$	$(10 - 12)^2 = 4$	$(18 - 12)^2 = 36$
$(6 - 12)^2 = 36$	$(11 - 12)^2 = 1$	$(22 - 12)^2 = 100$
$(8 - 12)^2 = 16$	$(12 - 12)^2 = 0$	$(27 - 12)^2 = 225$
$(9 - 12)^2 = 9$	$(12 - 12)^2 = 0$	$(49 - 12)^2 = 1369$

**Сума** = 2370; **Дисп** = 2370/26 = 91.15; **SD** =  $\sqrt{91.15} = 9.55$



# Нормальний розподіл



# Ціль 2

Обговорити приклади використання мір поширення захворювання в епідеміології





# Приклад: Використання діапазону в епідеміології

Діапазон зазвичай використовується при складанні звітів:

- інкубаційний період
- тривалість захворювання
- вік випадків

## Коротка характеристика окремих засобів біологічної війни

Agent	Incubation period (days)	Person to person transmission	Morbidity/mortality if untreated	Diagnosis
<i>Ba. anthracis</i>	1–5	No	High/high	Culture, serology, ELISA, PCR
<i>Y. pestis</i>	2–3	Yes	High/high	Culture, serology, ELISA, PCR
<i>Brucella sp.</i>	5–60	No	High/low	Culture, serology, ELISA, PCR
<i>Bu. mallei</i>	3–7	No	High/low	Culture, serology, ELISA, PCR
<i>Bu. pseudomallei</i>	3–7	No	High/low	Culture, serology, ELISA, PCR
Botulinum toxin	1–5	No	High/high	ELISA, mouse inoculation for toxin detection
Variola virus	7–17	Yes	High/high	Serology, virus isolation, ELISA, PCR
Viral hemorrhagic fever	4–21	Yes	High/high	Serology, virus isolation, ELISA, PCR

Source: Thavaselvam and Vijayaraghavan. *J Pharm Bioallied Sci.* 2010 Jul-Sep; 2(3): 179–188



# Приклад: Використання міжквартильного діапазону

Міжквартильний діапазон (IQR) використовується, коли наявні випадки можуть спотворити дані

## Changing Epidemiology of Measles in Africa

James L. Goodson,<sup>1</sup> Balcha G. Masresha,<sup>2</sup> Kathleen Wannemuehler,<sup>1</sup> Amra Uzicanin,<sup>1</sup> and

<sup>1</sup>Global Immunization Division, Centers for Disease Control and Prevention, Atlanta, Georgia; and <sup>2</sup>Immunisation and World Health Organization, Africa Regional Office, Brazzaville, Congo

**Background.** In Africa before the introduction of measles vaccination, measles children. To describe measles epidemiology in Africa since the start of accelerated measles we analyzed regional measles case-based surveillance data for 2002–2009.

**Methods.** Country-years were grouped by 10-year moving average of routine measles coverage (aMCV1). Age was log transformed, and pair-wise comparisons of means were made. A  $\chi^2$  test was used to assess association between coverage and age groups. Cumulative percent curves and percentiles of age, dot plots with Loess curve, and Spearman rank correlation coefficient were calculated.

**Results.** Of 180,284 suspected cases, 73,009 (41%) were confirmed as measles. Of these, the mean age was 79 months (median, 36 months; interquartile range, 16–96 months) and significantly younger in country-years with <50% aMCV1 than those with 50%–74% aMCV1 ( $P = .03$ ) and  $\geq 75%$  ( $P = .02$ ). With increasing coverage, there was a slight decrease in age in the 10th and 25th and moderate increase in age in the 50th, 75th, and 90th percentiles.

**Conclusions.** During 2002–2009, the median age of confirmed measles was 36 months. In countries with  $\geq 50%$  aMCV1 coverage compared with low-coverage countries, age shifted to older children and young adults; for infants, age decreased slightly with higher coverage.

...suspected cases, 73,009 (41%) were confirmed as measles (median age, 36 months; interquartile range, 16–96 months) and significantly younger in country-years with 50%–74% aMCV1 ( $P = .03$ ) and  $\geq 75%$  ( $P = .02$ ). With increasing coverage, there was a slight decrease in age in the 10th and 25th and moderate increase in age in the 50th, 75th, and 90th percentiles.



# Приклад: Використання стандартного відхилення в епідеміології

Стандартне відхилення часто використовуються разом із середнім значенням, щоб показати мінливість даних

Таблиця 1. Структура тіла і тяжкість захворювання серед ВІЛ-позитивних і ВІЛ-негативних осіб з/без туберкульозу

Characteristic [mean, (SD)] <sup>c</sup>	HIV positive with TB (n=31)		HIV negative with TB (n=32)	
	Men (n=10)	Women (n=21)	Men (n=18)	Women (n=14)
Age in yrs	30.9 (4.6)	29.2 (5.9)	26.0 (7.3)	26.3 (4.6)
BMI kg/m <sup>2</sup>	18.4 (1.7)	18.6 (3.0)	18.2 (2.0)	20.3 (4.3)
LMI in kg/m <sup>2</sup>	16.6 (1.3)	15.4 (0.9) <sup>b</sup>	16.6 (1.5)	16.1 (0.9)
FMI in kg/m <sup>2</sup>	1.8 (0.6)	3.3 (2.1)	1.6 (0.8)	4.6 (3.6) <sup>a</sup>
Severity TBscore	6.5 (1.5)	5.6 (2.6)	6.8 (2.3)	5.9 (2.4)

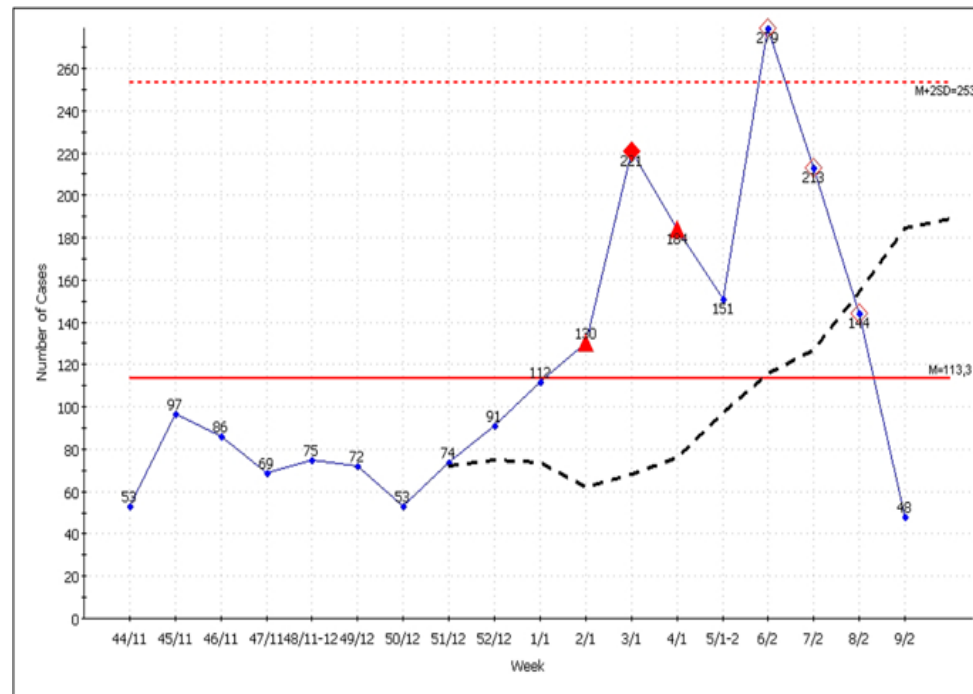
Source: Mupere et al. BMC Public Health 2012, 12:1050



# Приклад: Використання стандартного відхилення в епідеміології

Стандартне відхилення допомагає виявити незвичайне збільшення захворювань і може бути використане для встановлення граничних значень

**Спалах діареї після затоплення в Джакарті, Індонезія, січ–лют 2002**



← 2 SD (стандартне відхилення)

← Середнє значення



Cusum Flags: C1-Mild Sensitivity C2-Moderate Sensitivity C3-Ultra Sensitivity  
--- Avg Last 7 Intervals □=C1 ■=C1C3 ▲=C2C3 ◆=C3 ◆=C1C2C3

Source: Chretien et al, PLoSmedicine, Mar 2008 Vol 5(3).



Animal Health Matters.  
For Safe Food Solutions.



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,  
Education and Research EAER

**State Secretariat for Economic Affairs SECO**

Дякую