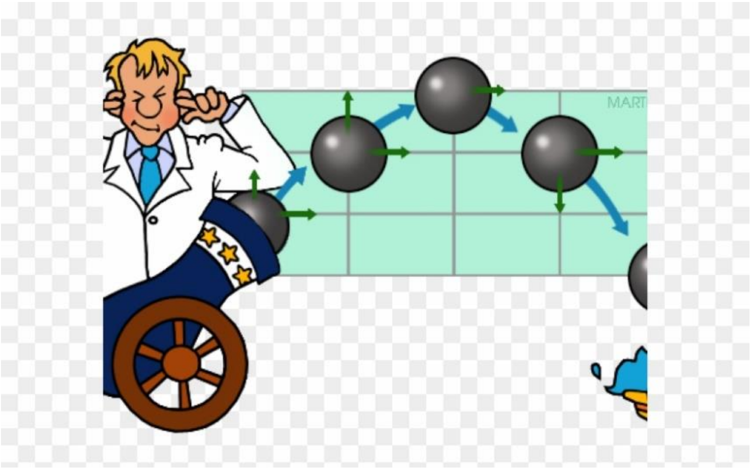
Laboratorium Podstaw Fizyki
zajęcia wprowadzające – kurs 15h

Krzysztof Gałkowski

- krzysztof.galkowski@pwr.edu.pl



Wszystkie ważne linki!

- Nasza strona http://lpf.wppt.pwr.edu.pl/index.php

Regulamin: http://lpf.wppt.pwr.edu.pl/informacje/regulamin.pdf

- http://lpf.wppt.pwr.edu.pl/informacje/uwagi_dot_przygotowywania_sprawozdan.pdf
- http://lpf.wppt.pwr.edu.pl/sprawozdania.php

- http://lpf.wppt.pwr.edu.pl/pomoce/Jaknarysowacwykres.pdf

- http://lpf.wppt.pwr.edu.pl/informacje/an_n.pdf niepewności pomiarowe

- http://lpf.wppt.pwr.edu.pl/informacje/rl.pdf (regresja liniowa - od 3. ćwiczeń)

Gotowość na indywidualne potrzeby

osoby, które ze względu na stan zdrowia, niepełnosprawność lub inne obiektywne przesłanki mogą mieć szczególne potrzeby związane ze sposobem realizacji zajęć, zaliczenia bądź przygotowaniem materiałów proszone są o zgłoszenie się na konsultacje, napisanie takiej informacji na prywatnym czacie, bądź napisanie e-maila w tej sprawie. Będę starał się, aby na moich zajęciach każdy miał <u>równe prawo do zdobycia wiedzy i rozliczenia się z niej.</u>

PWr podzieli się z prowadzącymi zajęcia informacją o szczeólnych potrzebach związanych z orzeczonymi niepełnosprawnościami, jednak ujawnienie innych czynników, jak np. ciąża, czy lęk przed wystąpieniami publicznymi, leży w gestii studentów

Podstawowe definicje

- Pomiar. Zespół czynności wykonywanych w celu ustalenia wartości wielkości fizycznej
- Niepewność pomiaru. Parametr określający ilościowo, jak wynik pomiaru może odbiegać od rzeczywistej wartości mierzonej wielkości fizycznej
- Niepewność standardowa pomiaru x. Oznaczana $u(x)$, wyróżniamy niepewności typu A i B
- Typ A: Określana poprzez statystyczną analizę serii pomiarów (przypadkowy rozkład wyników), np. odchylenie standardowe
- Typ B: Czynniki niestatystyczne, jak precyzja urządzeń pomiarowych oraz... osób przeprowadzających pomiary! Obecne praktycznie we wszystkich pomiarach!
- Niepewności złożone $u_c(x)$: uwzględniające współistnienie niepewności typów A i B.

Niektóre źródła niepewności pomiarowych

- Ograniczona precyzja urządzenia pomiarowego

- Niedoskonałość metody pomiarowej
-również “błędy grube” – błędy ludzkie prowadzące do absurdalnych wyników (np. pomyłka przy zapisie wartości)

- Nieprawidłowe wartości danych i materiałów referencyjnych

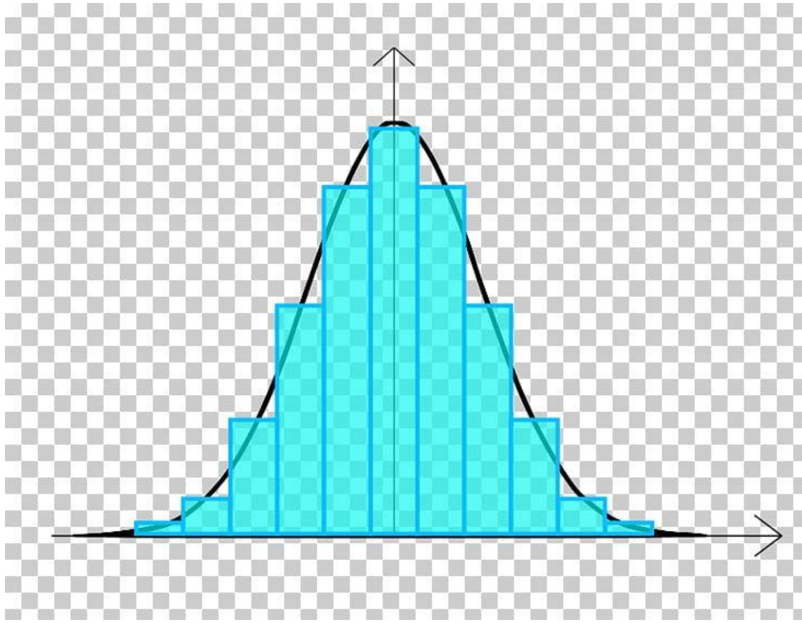
- Przybliżenia i założenia w procesie eksperymentalnym

- Rozrzut wyników pomiarów powtarzanych w nominalnie identycznych warunkach

- Niedostateczna wiedza nt. wpływu otoczenia na mierzoną wielkość

Więcej o niepewnościach standardowych typów A i B

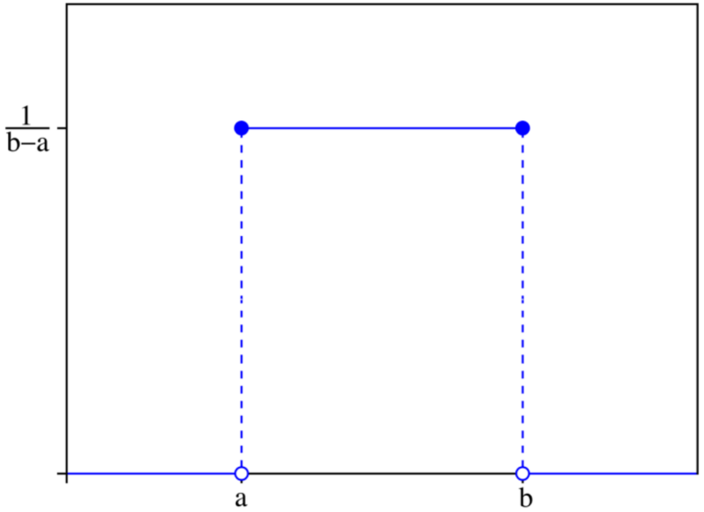
- **Typ A,** wiąże się ze statystycznym rozkładem wyników (serii) pomiarów
- Określane metodami statystycznymi – odchylenie standardowe, regresja liniowa
- Normalny rozkład prawdopodobieństwa (krzywa Gaussa)



- **Typ B,** zwykle szacujemy na podstawie znajomości błędów urządzeń pomiarowych ...oraz własnego, zdroworozsąkowego oszacowania!
- Czy urządzenie jest właściwie skalibrowane?

Czy nie należy uwzględnić niepewności związanej z „czynnikiem ludzkim”?

- Spodziewamy się jednostajnego rozkładu prawdopodobieństwa



Oszacowanie niepewności typu A (pomiar bezpośredni)

- Pomiar bezpośredni – mierzymy pojedynczą wielkość, jedną metodą pomiarową, np. masę za pomocą wagi, długość stołu za pomocą taśmy mierniczej

- Załóżmy, że wartość parametru *X* badamy wykonując *n* identycznych, bezpośrednich pomiarów x_1, x_2,…, x_n. Najbardziej prawdobodobną wartością *X* będzie więc średnia wszystkich pomiarów

$$x \equiv \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Spodziewamy się rozkłau normalnego (Gaussa) wyniku pomiarw, niepewności oszacujemy więc za pomocą odchylenia standardowego:

Niepewność pojedynczego pomiaru x_i:

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

oraz $\frac{1}{\sqrt{n}}$ razy mniejsza niepewność średniej $\bar{x}$:

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

Oszacowanie niepewności typu A (pomiar bezpośredni)

- "odmierzamy 1 sekundę" za pomocą stopera w smartfonie, 10 razy

	t_i [s]	$(t_i - \bar{t})^2$ [s²]
t_1	1,06	0,0028
t_2	1,09	0,0005
t_3	1,17	0,0032
t_4	1,02	0,0086
t_5	1,05	0,0040
t_6	1,18	0,0045
t_7	1,27	0,0246
t_8	1,19	0,0059
t_9	1,03	0,0069
t_{10}	1,07	0,0018
$\bar{t}$	**1,1130**	**0,0630**

$$\bar{t} = \frac{1}{10}\sum_{1}^{10} t_i$$ średnia

$$\sum_{1}^{10}\left(t_i - \bar{t}\right)^2$$

$$x \equiv \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad u(x) = \sqrt{s_{\bar{x}}^2} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

$$n = 10$$

$$u(t) = \sqrt{s_{\bar{t}}{}^2} = \sqrt{\frac{1}{10(10-1)}\sum_{1}^{10}(t_i - \bar{t})^2} = \sqrt{\frac{1}{90}0.0630} = 0{,}0264575 \text{ s}$$

Jak zapisać wynik i niepewność?

- "odmierzamy 1 sekundę" za pomocą stopera w smartfonie, 10 razy

$$\bar{t} = \frac{1}{10}\sum_{1}^{10} t_i = 1{,}1130\ \mathrm{s} \qquad u(t) = \sqrt{{s_{\bar{t}}}^{2}} = 0{,}0264575\ \mathrm{s}$$

Rozpoczynamy od niepewności, zaokrąglając do <u>dwóch cyfr znaczących</u>. Przyjmiemy konwencję <u>zaokrąglania niepewności zawsze w górę</u> – „wbrew" zasadom matematycznym

$$u(t) = 0{,}0264575\ s = 0{,}027\ s$$

Następnie, zaokrąglamy wynik do takiej samej liczby miejsc po przecinku, tym razem już według reguł matematycznych. Tym samym:

t = 1,113 s, *u(t)* = 0,027 s.

Więcej o zaokrąglaniu niepewności i wyników na kolejnych slajdach

Uwaga: Przyjęto, że w Polsce jako separatora części dziesiętnej używamy przecinka.

Oszacowanie niepewności typu A (pomiar bezpośredni)

t = 1,113 s, $u(t)$ = 0,027 s.

	t_i [s]	$\bar{t}$ from t_1 to t_i [s]
t_1	1,06	1,060
t_2	1,09	1,075
t_3	1,17	1,107
t_4	1,02	1,085
t_5	1,05	1,078
t_6	1,18	1,095
t_7	1,27	1,120
t_8	1,19	1,129
t_9	1,03	1,118
t_{10}	1,07	1,113

Może zabrzmi to banalnie, im większa próbka danych tym średnia lepiej reprezentuje rzeczywisty wynik ...ale zaraz, a co się tutaj wydarzyło ?!

„Nasza" długość jednej sekundy jest istotnie wyższa od „rzeczywistej", „oczekiwanej" 1,000 s , a różnica między nimi jest kilkukrotnie większa od niepwewności. Czy przeprowadzony rachunek niepewności ma w ogóle sens?

Tak. Bez wątpienia obserwujemy rozrzut statystyczny punktów pomiarowych, analiza niepewności typu A jest zasadna – niepewność średniej określona przez odchylenie standardowe daje nam informację o rozrzucie pojedynczych wyników pomiarów.

Jednak to nie wszystko. Nie uwzględniliśmy precyzji stopera, czasu reakcji eksperymentatora, a ponadto oczywistego *błedu systematycznego* – osoba wykonująca pomiar miała wyraźną tendencję do zatrzymywania pomiaru zbyt późno (w serii powinny pojawić się wyniki zarówno powyżej, jak i poniżej wartości oczekiwanej).
Przejdźmy więc do niepewności typu B!

Pytanie do samodzielnego rozważenia: Jak zwiększyć dokładność takiego pomiaru?

Niepewności typu B (w pomiarze bezpośrednim)

- Związane z błędem („błędem maksymalnym”) przyrządu Δ*x*. Jeśli wynik pomiaru to x, wtedy rzeczywista wartość mierzonej wielkości mieści się w przedziale $\langle x - \Delta x, x + \Delta x \rangle$, z prawdopodobieństwem opisywanym przez rozkład jednostajny
- Wtedy niepewność *u(x)* definiujemy jako:

$$u(x) = \frac{\Delta x}{\sqrt{3}} = \sqrt{\frac{(\Delta x)^2}{3}}$$

- Często należy uwzględnić też błąd eksperymentatora Δx_e, którego dokładność może być istotnie gorsza od dokładności urządzenia pomiarowego. Przykład – mierzymy czas stoperem o rozdzielczości 0.01 s, nasz czas reakcji to przynajmniej 0.1 s...
- Jeśli 2 (lub więcej...) źródła niepewności:

$$u(x) = \sqrt{\frac{(\Delta x)^2}{3} + \frac{(\Delta x_e)^2}{3} + \ldots}$$

- Niepwność złożona – gdy jednocześnie uwzględniamy niepewności typów A i B:

$$u(x) = \sqrt{s_{\bar{x}}^2 + \frac{(\Delta x)^2}{3} + \frac{(\Delta x_e)^2}{3}}\,.$$

Niepewności typu B – określanie błędu przyrządu pomiarowego

- Urządzenia „analogowe”: np. taśma miernicza, suwmiarka: Δx określa się jako połowę podziałki

- Urządenia „cyfrowe”: $\Delta x = c_1 x + c_2 z$. Stałe c_1, c_2 podawane są przez producenta. Pierwsze wyrażenie – składnik proporcjonalny do odczytu (zmierzonej wartości) *x*. Zwykle podawana w procentach, stała c_1 nazywana jest *klasą* przyrządu pomiarowego. Drugie wyrażenie to składnik stały związany z rozdzielczością przyrządu *z*.

- **Przykład:** Dokonujemy pomiaru napięcia U = 89,71 V. Instrukcja multimetru określa maksymalny błąd na ±(0,5 % rdg + 1 dgt). Jak to odczytać? c_1 = 0.5 %, c_2 = 1; „rdg” to „odczyt”, a „dgt” to rozdzielczość – w tym wypadku wynosi 0,01 V (odczytujemy z dokładnością do setnych części volta)

- ΔU = 89,71*0,5 % + 1*0,01= 0,45855 V. Niepewność $u(U) = \Delta U / \sqrt{3}$ = 0,26475 V = 0,27 V.

Niepewność rozszerzona pomiaru

- Jakie jest prawdopodobieństwo, że rzeczywista wartość mierzonej wielkości *X* leży w zakresie wyznaczonym przez niepewność standardową *u(x)*, tj. $\langle \bar{x} - u(x), \bar{x} + u(x) \rangle$?
- Dla typu A (Gauss) 68%; typ B (rozkład jednostajny): 58%
- Niepewność rozszerzona *U(x)*: niepewność standardowa pomnożona przez <u>współczynnik rozszerzenia</u> *k*

$$U(x) = ku(x)$$

aby zwiększyć prawdopodobieństwo, że rzeczywista wartość mierzonego parametru leży w podawanym zakresie niepewności.

- Najcześciej *k* = 2, wtedy prawdopodobieństwo 95 % dla typu A (>99 % dla *k* = 3)* oraz 100 % dla typu B (osiągane już √3, wtedy *U*(*x*) równe jest maksymalnemu błedowi urządzenia pomiarowego.
- N. rozszerzonych używa się w zastosowaniach, gdzie wymagane jest wysokie prawdopodobieństwo (lub *poziom ufności*, również wyrażany w %), że rzeczywista wartość badanego patametru *istotnie* znajduje się w podawanym zakresie wartości. Metrologia, kalibracja przyrządów laboratoryjnych, przemysł, budownictwo

*podane wartości *k* dla niepewności typu A zakładają duże wartości *n.* Aby osiągnąć poziom ufności 95% dla *n* = 2, *k* ≈ 13, natomiast *k* ≈ 3 dla *n* = 5.

Zaokrąglanie i zapis wyniku z niepewnością pomiaru (I)

Przykład: x = 127,34287 m, u_x = 0,87124 m. Po zaokrągleniu, u_x = 0,88 m, więc wynik zaokrąglamy do x = 127,34 m:

(1) $x = 127{,}34(88)\ \mathrm{m}$

Symbole oznaczające parametry podajemy *kursywą*

spacje

natomiast jednostki – „normalnym” tekstem

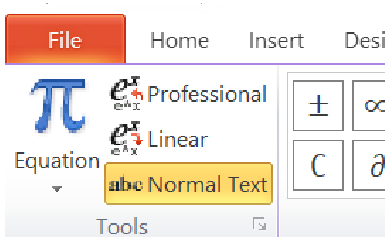
(2) $x = (127{,}34 \pm 0{,}88)\ \mathrm{m}$

Konwencja zapisu, kursywy i spacje, <u>są istotne</u>!

Zapis (1) poprawny dla niepewności standardowych; dwie cyfry w nawiasie oznaczające niepewność odpowiadają dwum ostatnim cyfrom w zapisie wyniku. Uwaga: do nawiasu nie wprowadzamy separatora części dziesiętnych
Zapis (2), “plus-minus” typowy i poprawny dla niepewności rozszerzonych.
We współczesnej komunikacji naukowej zasada ta nie jest zawsze ściśle przestrzegana, niepewności są poniekąd „z definicji” uznawane za rozszerzonei zapis “±” spotykamy jako standardowy w wielu czasopismach. My jednak będziemy rozróżniać obydwa zapisy!

Porada: Narzędzie Microsoft Equation automatycznie zadba o odpowiednie spacje w zapisie matematycznym, należy jedynie pamiętać o wyborze “normalnego tekstu” gdy zapisujemy jednostkę.
Przykładowy zapis za pomocą MEq:

$$x = 127{,}34(88)\ \mathrm{m}$$



Zaokrąglanie i zapis wyniku z niepewnością pomiaru (II)

Kilka przykładów zaokrąglania i zapisu (dla standardowej niepewności pomiaru – notacja z nawiasami)

Uncertainity	Quantity	Correct notation
u_T = 3,15 K ≈ 3,2 K	T = 293,12 K ≈ 293,1 K	T = 293,1(32) K
u_t = 0,000015 s	t = 1,123 s	t = 1,123000(15) s
u_d = 2230 kg ≈ 2300 kg	d = 78327 kg ≈ 78300 kg	d = 78300(2300) kg
u_r = 0,0000025 mm ≈ $2{,}5\cdot10^{-9}$ m	r = 0,0000450 mm ≈ $4{,}5\cdot10^{-8}$ m	$r = 4{,}50(25)\cdot10^{-8}$ m

Konsturuując tabelę, zwykle podajemy najpier wartością zmierzonej wielkości, następnie niepewność; tutaj dokonano wyjątkowego przestawienia w celu ułatwienia czytelności ze względu na kolejność zaokrąglania (najpierw niepewność) – jest to ćwiczenie, a nie sprawozdanie

Obliczenia niepewności łączonej:

- Za pomocą analogowej suwmiarki o podziałce 0,1mm, 13 razy zmierzono średnicę otworu na śrubę M6 w ławie optycznej (tj. z nieskończonej liczby możliwych przyłożeń suwmiarki w celu pomiaru średnicy, wybrano 13). Wyniki to: 5,9, 5,9, 6,0, 6,1, 5,9, 6,1, 5,8, 6,0, 6,0, 5,9, 6,2, 5,9 oraz 6,0 mm.

  - Ile wynosi wynik pomiaru średnicy śruby wraz z niepewnością?

  - Staraj się samodzielnie odpowiadać na pytania, na podstawie wiedzy wyniesionej z tej prezentacji. Rozwiązanie na kolejnych slajdach

  - Na początek – jaki jest wynik pomiaru?

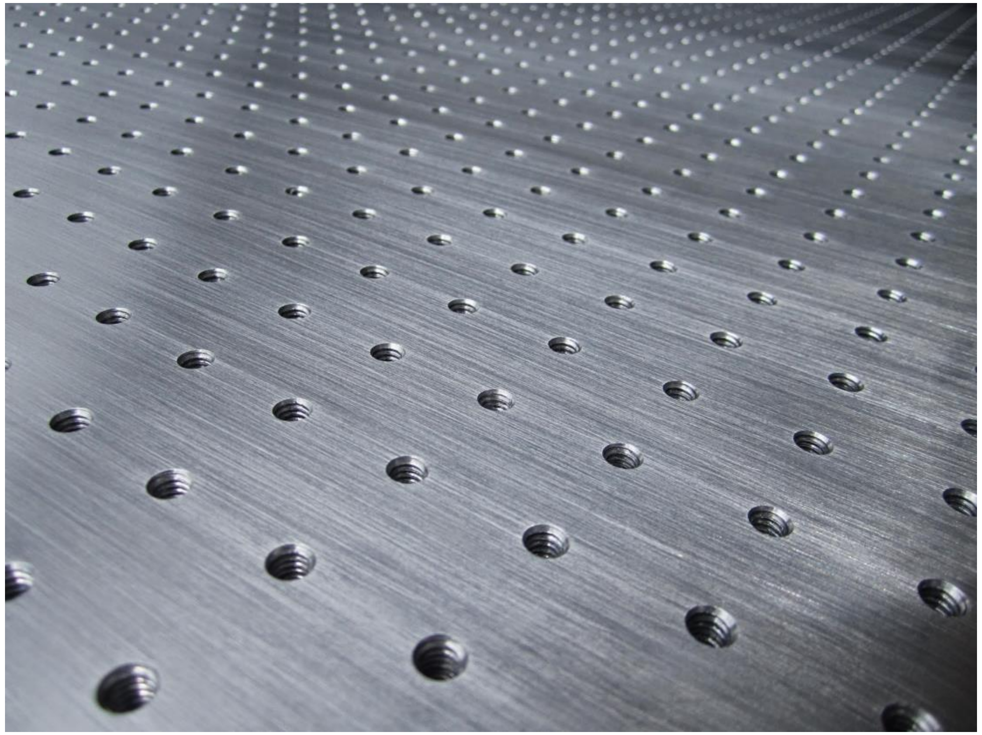


Photo: baselabtools.com

Obliczenia niepewności łączonej:

- Za pomocą analogowej suwmiarki o podziałce 0,1mm, 13 razy zmierzono średnicę otworu na śrubę M6 w ławie optycznej (tj. z nieskończonej liczby możliwych przyłożeń suwmiarki w celu pomiaru średnicy, wybrano 13). Wyniki to: 5,9, 5,9, 6,0, 6,1, 5,9, 6,1, 5,8, 6,0, 6,0, 5,9, 6,2, 5,9 oraz 6,0 mm.

- Wynik pomiaru najlepiej reprezentuje średnia: $d \equiv \bar{d} = \frac{1}{13}\sum_{1}^{13} d_i = 5{,}9769231\ \mathrm{mm}$

- Czy występują niepewności typu B?

Obliczenia niepewności łączonej:

- Za pomocą analogowej suwmiarki o podziałce 0,1mm, 13 razy zmierzono średnicę otworu na śrubę M6 w ławie optycznej (tj. z nieskończonej liczby możliwych przyłożeń suwmiarki w celu pomiaru średnicy, wybrano 13). Wyniki to: 5,9, 5,9, 6,0, 6,1, 5,9, 6,1, 5,8, 6,0, 6,0, 5,9, 6,2, 5,9 oraz 6,0 mm.

- Wynik najlepiej reprezentuje średnia: $d \equiv \bar{d} = \frac{1}{13}\sum_{1}^{13} d_i = 5{,}9769231\ \mathrm{mm}$
- Czy występują niepewności typu B?

- Tak, związane z błędem urządzenia pomiarowego:
- Maksymalny błąd przyrządu (połowa podziałki): Δ*d* = 0,1/2 = 0,05 mm
- Niepewność typu B: $u_B(d) = \frac{\Delta d}{\sqrt{3}} = \frac{0{,}05}{\sqrt{3}} = 0{,}02887\ \mathrm{mm}$

- Czy występują niepewności typu A?

Obliczenia niepewności łączonej:

- Za pomocą analogowej suwmiarki o podziałce 0,1mm, 13 razy zmierzono średnicę otworu na śrubę M6 w ławie optycznej (tj. z nieskończonej liczby możliwych przyłożeń suwmiarki w celu pomiaru średnicy, wybrano 13). Wyniki to: 5,9, 5,9, 6,0, 6,1, 5,9, 6,1, 5,8, 6,0, 6,0, 5,9, 6,2, 5,9 oraz 6,0 mm.

  - Wynik najlepiej reprezentuje średnia: $d \equiv \bar{d} = \frac{1}{13}\sum_{1}^{13} d_i = 5{,}9769231\ \mathrm{mm}$
  - Maksymalny błąd przyrządu (połowa podziałki): Δd = 0,1/2=0,05 mm
  - Niepewność typu B: $u_B(d) = \frac{\Delta d}{\sqrt{3}} = \frac{0{,}05}{\sqrt{3}} = 0{,}02887\ \mathrm{mm}$

  - Czy występują niepewności typu A? Tak, obserwujemy statystyczny rozrzut wyników. Obliczmy niepewność średniej używając odchylenia standardowego:

$$u_A(d) = \sqrt{\frac{1}{13(13-1)}(d_i - 5{,}9769231)^2} = 0{,}0302846\ \mathrm{mm}$$

  - Ile wynosi całkowita niepewność pomiaru?

Obliczenia niepewności łączonej:

- Za pomocą analogowej suwmiarki o podziałce 0,1mm, 13 razy zmierzono średnicę otworu na śrubę M6 w ławie optycznej (tj. z nieskończonej liczby możliwych przyłożeń suwmiarki w celu pomiaru średnicy, wybrano 13). Wyniki to: 5,9, 5,9, 6,0, 6,1, 5,9, 6,1, 5,8, 6,0, 6,0, 5,9, 6,2, 5,9 oraz 6,0 mm.

  - Wynik najlepiej reprezentuje średnia: $d \equiv \bar{d} = \frac{1}{13}\sum_{1}^{13} d_i = 5{,}9769231\ \mathrm{mm}$
  - Maksymalny błąd przyrządu (połowa podziałki): Δ*d* = 0,1/2=0,05 mm
  - Niepewność typu B: $u_B(d) = \frac{\Delta d}{\sqrt{3}} = \frac{0{,}05}{\sqrt{3}} = 0{,}02887\ \mathrm{mm}$
  - Niepewność typu A: $u_A(d) = \sqrt{\frac{1}{13(13-1)}(d_i - 5{,}9769231)^2} = 0{,}0302846\ \mathrm{mm}$
  - Ile wynosi całkowita niepewność pomiaru?
  - Niepewność łączona: $u_c(d) = \sqrt{{s_{\bar{d}}}^2 + \frac{(\Delta d)^2}{3}} = \sqrt{0{,}0302846^2 + 0{,}02887^2} = 0{,}041841 = 0{,}042\ \mathrm{mm}$

  - *d* = 5,977(42) mm

Niepewność pomiarów pośrednich

- W niektórych przypadkach, obliczamy interesującą nas wielkość za pomocą pomiaru innych wielkości (pomiar pośredni)

- Przykład. Chcemy wyznaczyć szybkość. Musimy więc jednocześnie zmierzyć przemieszczenie *s* oraz czas *t*, w którym to przemieszczenie nastąpiło. Następnie, obliczamy szybkość $v = s / t$.

- Ogólnie, jeśli wielkość z jest funkcją k parametrów $x_1, x_2, \ldots, x_k$, możemy zapisać $z = z(x_1, x_2, \ldots, x_k)$

- Niepewność pomiarów pośrednich otrzymujemy za pomocą tzw. prawa propagacji niepewności:

$$u_c(z) = \sqrt{\sum_{j=1}^{k} \left(\frac{\partial f(x_j)}{\partial x_j} \right)^2 u^2(x_j)} .$$

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k) .$$

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k).$$

Niepewność pomiarów pośrednich

$$u_c(z) = \sqrt{\sum_{j=1}^{k} \left(\frac{\partial f(x_j)}{\partial x_j} \right)^2 u^2(x_j)}.$$

Rozważmy szybkość *v* jako funkcję dwóch zmiennych *s* i *t*: $v(s,t) = \frac{s}{t}$

Wtedy niepewność *v* jako obliczonej pośrednio z pomiarów *s* i *t*: $u_c(v) = \sqrt{\left(\frac{\partial v(s)}{\partial s}\right)^2 {u_s}^2 + \left(\frac{\partial v(t)}{\partial t}\right)^2 {u_t}^2}$

Gdzie u_s, u_t to niepewności odpowiednio *s* i *t*. Obliczamy pochodne cząstkowe:

$$\frac{\partial v(s)}{\partial s} = \frac{1}{t}\frac{\partial}{\partial s} s = \frac{1}{t} * 1$$

$$\frac{\partial v(t)}{\partial t} = s \frac{\partial}{\partial t}\frac{1}{t} = s * \left(-\frac{1}{t^2}\right)$$

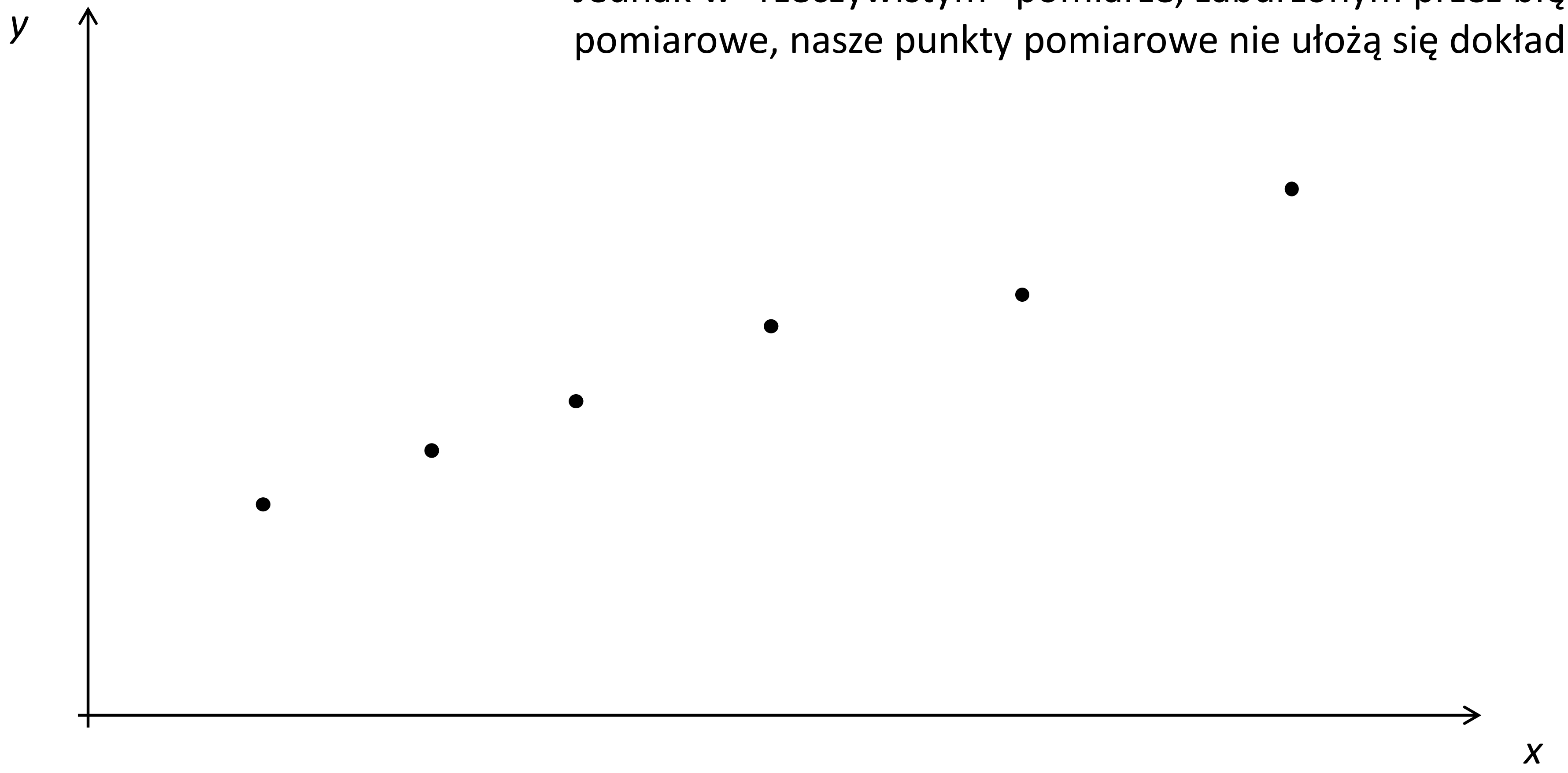
Ostatecznie:

$$u_c(v) = \sqrt{\left(\frac{1}{t}\right)^2 {u_s}^2 + \left(-\frac{s}{t^2}\right)^2 {u_t}^2}$$

Regresja liniowa

- Rozważmy zestaw następujący danych: pary parametrów x_i, y_i zmierzone n-razy. Wiemy, że pomiędzy wielkościami x,y powinna zachodzić zależność liniowa $y = Ax + B$.

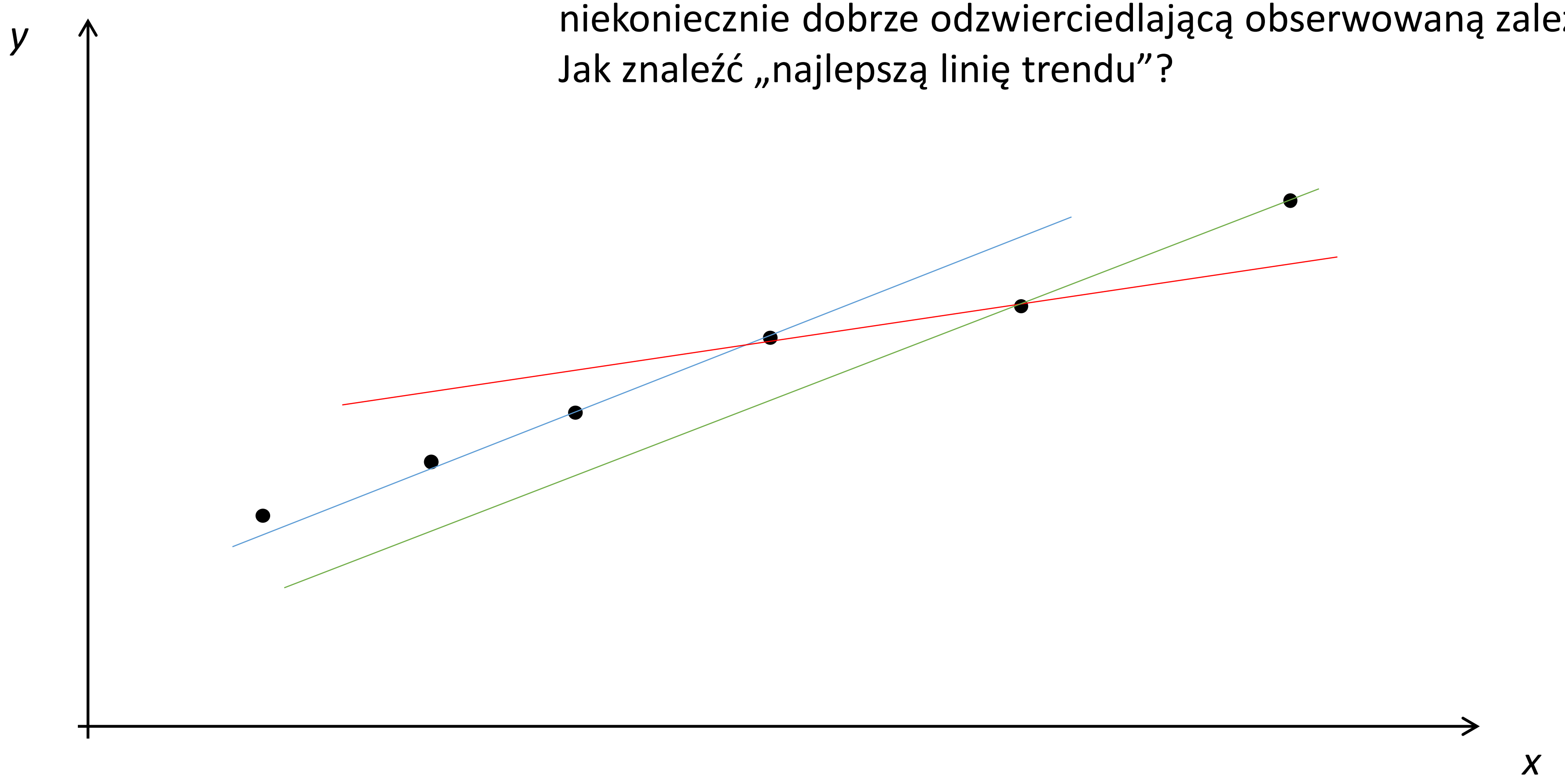
Jednak w "rzeczywistym" pomiarze, zaburzonym przez błędy/niepewności pomiarowe, nasze punkty pomiarowe nie ułożą się dokładnie na linii prostej…



Regresja liniowa

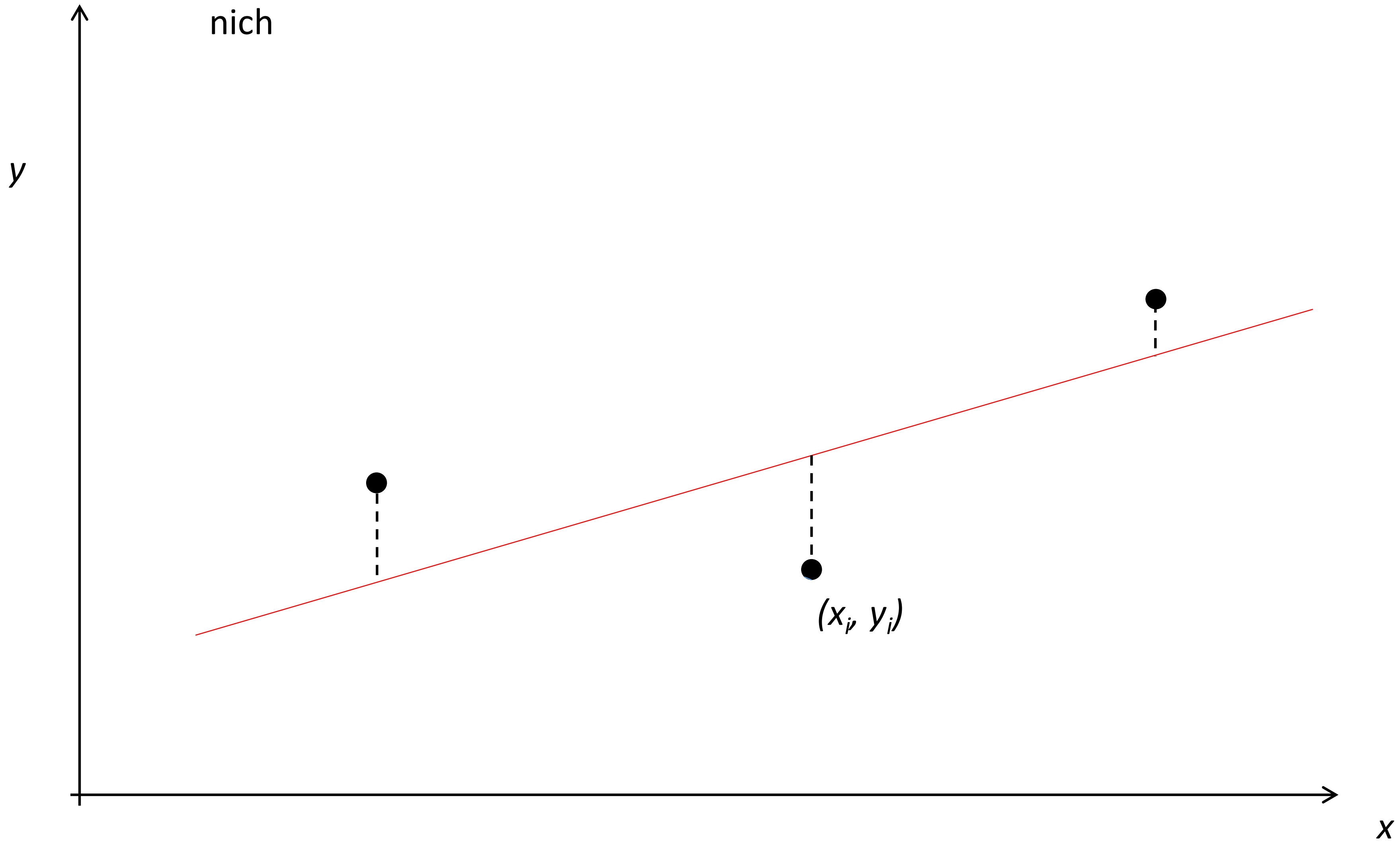
- Rozważmy zestaw następujący danych: pary parametrów x_i, y_i zmierzone *n*-razy. Wiemy, że pomiędzy wielkościami x,y powinna zachodzić zależność liniowa $y = Ax + B$

… więc tylko każde dwa wybrane punkty będziemy mogli połączyć linią, niekoniecznie dobrze odzwierciedlającą obserwowaną zależność
Jak znaleźć „najlepszą linię trendu”?



Regresja liniowa

Linia najlepiej reprezentująca trend nie musi przechodzić przez żaden z punktów pomiarowych! Wystarczy, że będzie „dostatecznie blisko” każdego z nich

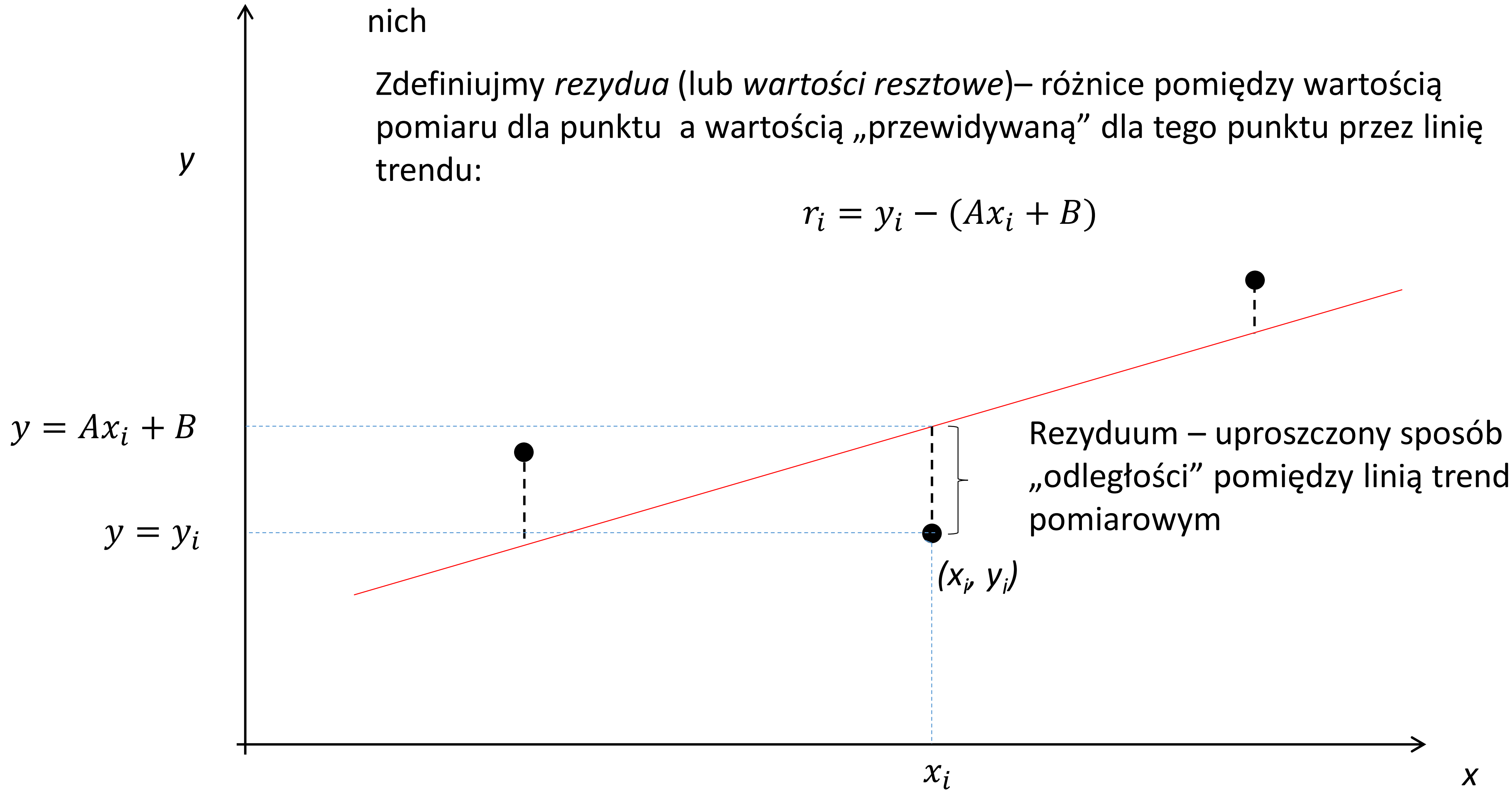


Regresja liniowa

Linia najlepiej reprezentująca trend nie musi przechodzić przez żaden z punktów pomiarowych! Wystarczy, że będzie „dostatecznie blisko” każdego z nich

Zdefiniujmy *rezydua* (lub *wartości resztowe*)– różnice pomiędzy wartością pomiaru dla punktu a wartością „przewidywaną” dla tego punktu przez linię trendu:

$$r_i = y_i - (Ax_i + B)$$



Rezyduum – uproszczony sposób zdefiniowania „odległości” pomiędzy linią trendu a punktem pomiarowym

Regresja liniowa

$$r_i = y_i - (Ax_i + B)$$

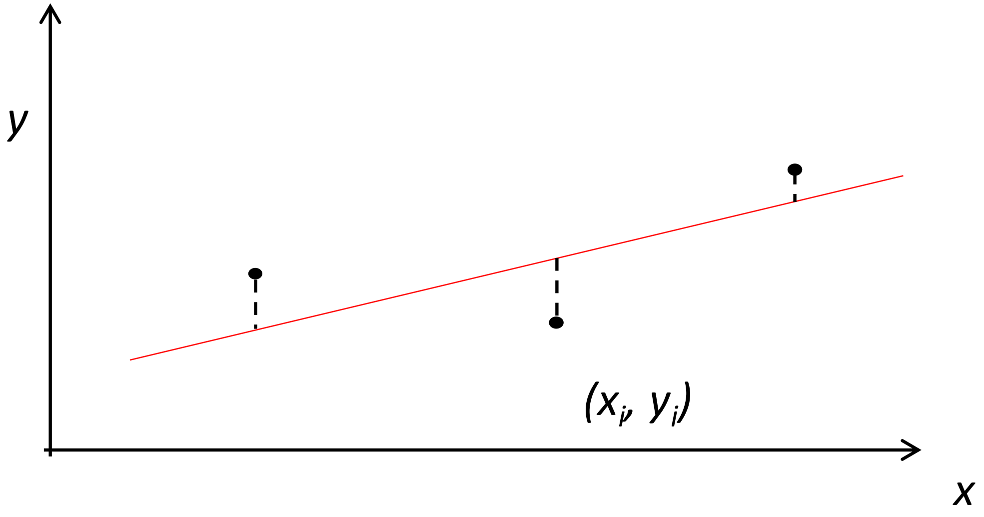
Chcemy „kolektywnie" dopasować linię trendu do wszystkich punktów pomiarowych – tak, aby suma „odległości" wszystkich n punktów (x_i, y_i) od linii trendu była jak najmniejsza. Rezydua mogą jednak przyjmować wartości ujemne lub dodatnie, aby więc aby operować na sumie „odległości", a nie matematycznej, zsumujmy ich kwadraty:

$$S^2 = \sum_{i}^{n} [y_i - (Ax_i + B)\,]^2$$

Aby znaleźć najmniejszą wartość S^2, szukamy pary paremetrów *A,B* spełniającej warunki:

$$\frac{\partial S^2}{\partial B} = 0 \text{ and } \frac{\partial S^2}{\partial A} = 0\,.$$

I tak docieramy do podstaw *metody najmniejszych kwadratów*, najpopularniejszej metody regresji liniowej



Więcej tutaj:
http://lpf.wppt.pwr.edu.pl/informacje/rl.pdf
(m.in. rozwiązania dla *A* i *B* jednak nie trzeba znać tych wyrażeń na pamięć)

Regresja liniowa

$$r_i = y_i - (Ax_i + B)$$

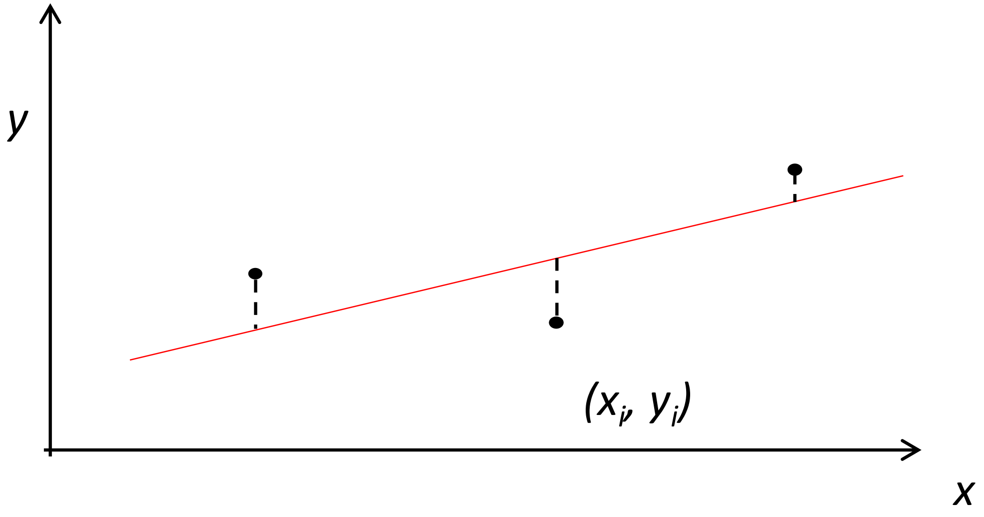
Chcemy „kolektywnie” dopasować linię trendu do wszystkich punktów pomiarowych – tak, aby suma „odległości” wszystkich *n* punktów (x_i, y_i) od linii trendu była jak najmniejsza. Rezydua mogą przyjmować wartości ujemne lub dodatnie, aby więc aby operować na sumie „odległości”, nie matematycznej, zsumujmy ich kwadraty:

$$S^2 = \sum_i^n [y_i - (Ax_i + B)\,]^2$$

Aby znaleźć najmniejszą wartość S^2, szukamy pary paremetrów *A,B* spełniającej warunki:

$$\frac{\partial S^2}{\partial B} = 0 \text{ and } \frac{\partial S^2}{\partial A} = 0\,.$$

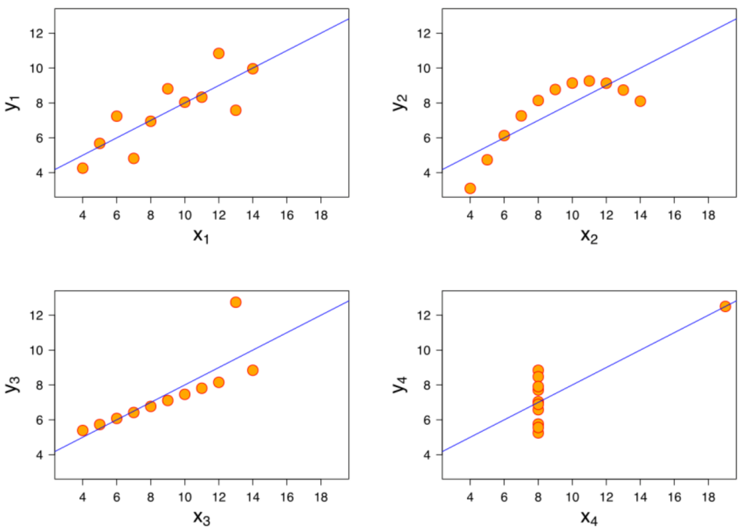
I tak docieramy do podstaw *metody najmniejszych kwadratów*, najpopularniejszej metody regresji liniowej



Regresja liniowa jest metodą statystyczną, zwracającą niepewności typu A. Niepewności typu B obliczone dla poszczególnych punktów nie zostaną uwzględnione przy obliczaniu parametrów *A,B* oraz ich niepewności. Dopasowania z uwzględnieniem niepewności łączonych możemy poznać na dodatkowych zajęciach.

Regresja liniowa

Pominięcie niepewności typu B to jedna z kilku słabości metody najmniejszych kwadratów



https://pl.wikipedia.org/wiki/Kwartet_Anscombe'a

Tutaj, kwartet Anscombe'a pokazuje jak bardzo metoda najmniejszych kwadratów jest czuła na tzw. punkty (obserwacje) odstajace Cztery graficznie różne zestawy 11 punktów mają niemal identyczny opis statystyczny

Dane doświadczalne zawsze analizujemy z należytą uwagą i „świadomie"!!!

Regresja linowa w Excelu:

Zaznacz 4 komórki, użyj funkcji

REGLIN.P(**znane_y**, [***znane_x***], [**stała**], [**statystyka**])

(w wersji angielskiej – LINEST())

[**stała**] – określa stałą B w $Ax + B$, PRAWDA jeśli zakładamy $B \neq 0$, FAŁSZ jeśli $B = 0$

[**statystyka**] – PRAWDA, jeśli chcemy dostać niepewności parametrów A i B

Wciśnij Ctrl + Shift + Enter aby otrzymać tabelę 2x2 :

Wartość A	**Wartość B**
Niepewność A	**Niepewność B**