Komputerowa analiza danych doświadczalnych

Wykład 12
12.05.2021

dr inż. Łukasz Graczykowski
lukasz.graczykowski@pw.edu.pl

Semestr letni 2020/2021

Weryfikacja hipotez statystycznych - ponownie

Metoda najmniejszych kwadratów

Weryfikacja hipotez statystycznych

- **Przykład:** rozważamy zmienną losową X opisaną standardowym rozkładem Gaussa (średnia 0, odchylenie 1). Pobieramy 10-elementową próbę, uzyskaliśmy średnią arytmetyczną: $\bar{X}=0{,}5$
- Jak na podstawie tej jednej realizacji próby (np. wyniku eksperymentu) możemy stwierdzić, czy pochodzi ona z takiej populacji? Innymi słowy, naszą **hipotezą** jest: **próba losowa pochodzi z rozkładu Gaussa o średniej 0 i odchyleniu 1**
- Procedura weryfikacji hipotezy nazywana jest **testem statystycznym**
- **Jeżeli hipoteza jest słuszna (nasze założenie)** to wartość średnia (będąca również zmienną losową) $\bar{X}$ ma rozkład normalny ze średnią 0 i odchyleniem std. $1/\sqrt{10}$

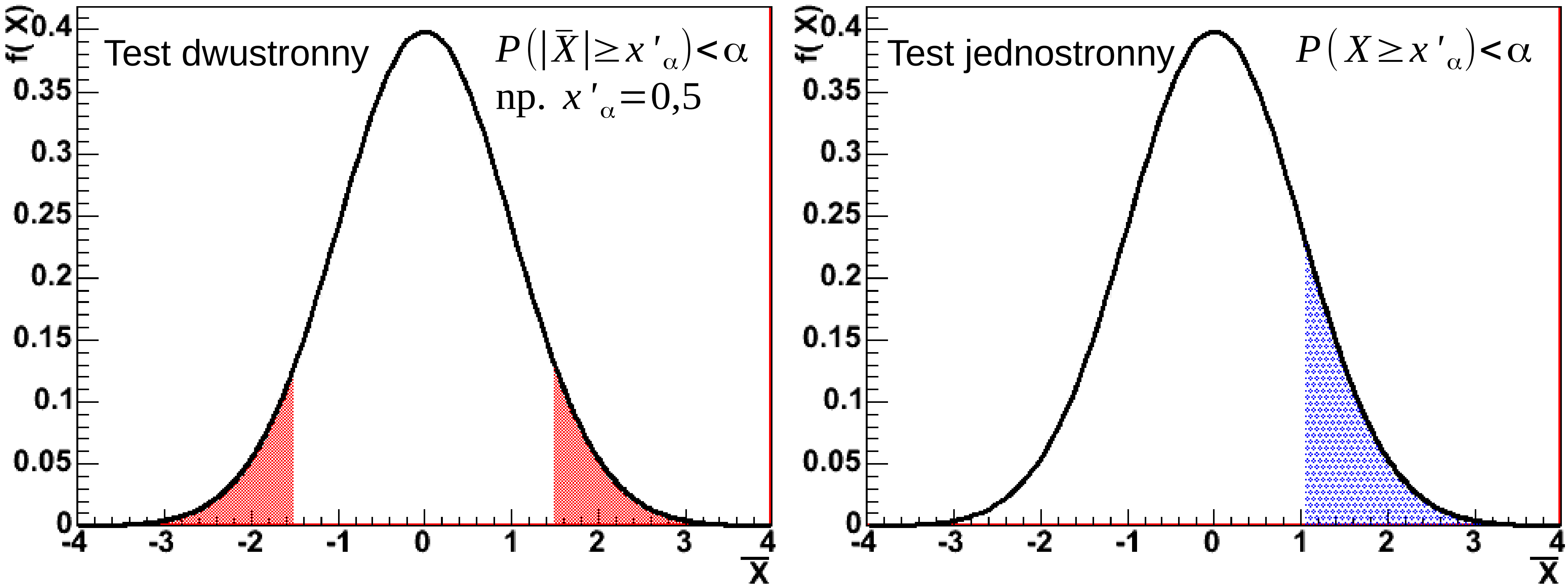
$$\sigma^2(\bar{X})=\frac{1}{n}\sigma^2(X)=\frac{1}{10}\cdot 1\Rightarrow\sqrt{\sigma^2(\bar{X})}=\frac{1}{\sqrt{10}}$$

Weryfikacja hipotez statystycznych

- Jak na podstawie **konkretnej realizacji próby** sprawdzić, czy założona hipoteza jest prawdziwa?
  - **I:** musimy ustalić pewną wartość prawdopodobieństwa α (zwanego **poziomem istotności**,z reguły mała wartość, np. 0,01, albo 0,03, czy 0,05)
  - **II:** pytamy, czy prawdopodobieństwo zaobserwowania określonych wartości próby jest mniejsze niż α: $P(|\bar{X}| \geq 0{,}5) < \alpha$
  - **nierówność <u>spełniona</u>** – jest mało prawdopodobne, aby próba pochodziła z rozkładu określonego przez testowaną hipotezę → **możemy ją odrzucić**
  - **prawdopodobieństwo zaobserwowania tego, że $|\bar{X}|$ jest duże, jest <u>bardzo małe</u>, ale takie nam się trafiło – więc prawdopodobnie (z prawdodobieństwem $1\text{-}\alpha$) nasza hipoteza nie jest słuszna**
  - **III:** jeśli prawdopodobieństwo jest mniejsze niż przyjęta wartość prawdopodobieństwa (poziom istotności) α, odrzucamy hipotezę na zadanym poziomie istotności

Weryfikacja hipotez statystycznych

Rozkład wartości średniej $\bar{X}$



- Jeśli (w naszym przykładzie) wartość średnia znajduje się w zaznaczonym obszarze (nazywamy go **obszarem krytycznym**), to hipotezę odrzucamy
  - jeśli oczekujemy rozkładu normalnego o średniej 0 i małym odchyleniu (np. 10), a z próby losowej (konkretny eksperyment) mamy średnią 1000, to lądujemy w **"ogonie"** rozkładu średniej i na podstawie <u>tej konkretnej próby</u> **odrzucamy** hipotezę (**ale na podstawie innej próby moglibyśmy zaakceptować**)

Weryfikacja hipotez statystycznych

- W ogólnym przypadku używamy innych wielkości niż średnia:
  - definiujemy jakąś (wygodną dla nas) statystykę testową T (np. różnicę między wynikiem eksperymentu a krzywą teoretyczną)
  - ustalamy poziom istotności α
  - wyznaczamy taki zbiór U, który określa obszar zmienności statystyki testowej T, taki że prawdopodobieństwo znalezienia się w nim jest ograniczone wartością α: $P(T \in U) = \alpha$
  - z pobranej próby wyznaczamy konkretną wartość statystyki testowej T': jeżeli znajduje się ona **wewnątrz** obszaru krytycznego U, **odrzucamy hipotezę** (mówimy: krzywa teoretyczna nie opisuje wyniku eksperymentu), czyli odrzucamy hipotezę, jeżeli $T' \in U$

Test dobroci χ^2 dopasowania

Test χ^2 dobroci dopasowania

- Mamy N pomiarów g_i, $i=1, 2, \ldots, N$ oraz ich niepewności σ_i
- Wartości $f_i,\ i=1,2,\ldots,N$ określają nam prawdziwy rozkład danej wielkości mierzonej (**np. znaleziony poprzez estymację**)
- **Dla każdego pomiaru liczymy wielkość u_i:** $u_i = \frac{g_i - f_i}{\sigma_i},\ i=1,2,\ldots,N$
- Jeśli nasza teoria (wartości f_i) jest prawdziwa, to rozkłady różnic u_i mają postać standardowego rozkładu normalnego **– nasza hipoteza**
- Jeśli tak, to rozkład χ^2 o N stopniach swobody będzie miała wielkość:

$$T = \sum_{i=1}^{N} u_i^2 = \sum_{i=1}^{N} \left(\frac{g_i - f_i}{\sigma_i} \right)^2$$

- **(Subiektywnie)** oczekujemy małej wartości wielkości T
- Gdy hipoteza jest **fałszywa**, wówczas poszczególne różnice u_i przyjmują duże wartości (wartość T jest duża)
- Jak określić granicę zmienności T? Można zauważyć, że granica ta jest określona **kwantylem** $\chi^2_{1-\alpha}$, czyli: $\boxed{P(T > \chi^2_{1-\alpha}) = \alpha}$

$$\boxed{\begin{aligned} F(x_q) = P(X \le x_q) = q \\ P(X > x_q) = 1 - q \end{aligned}}$$

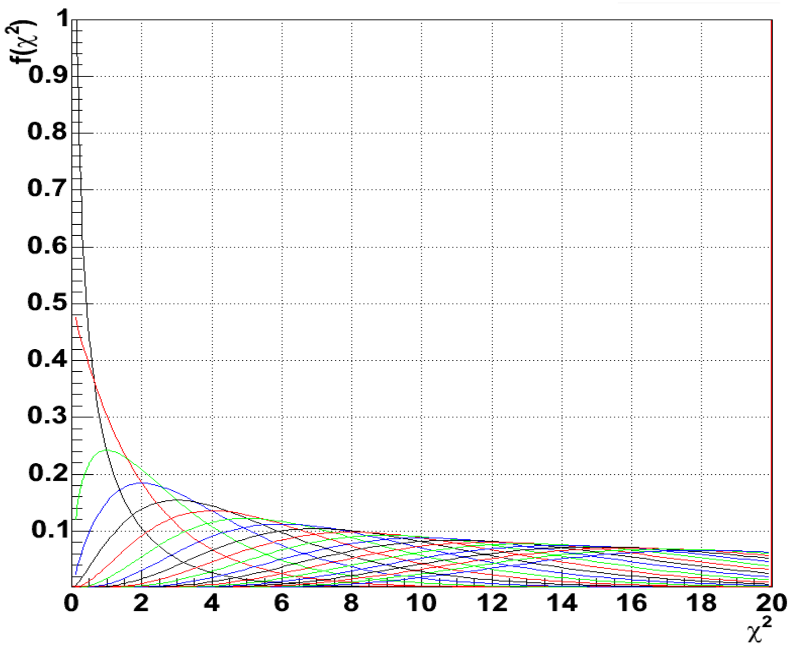
Test χ^2 dobroci dopasowania

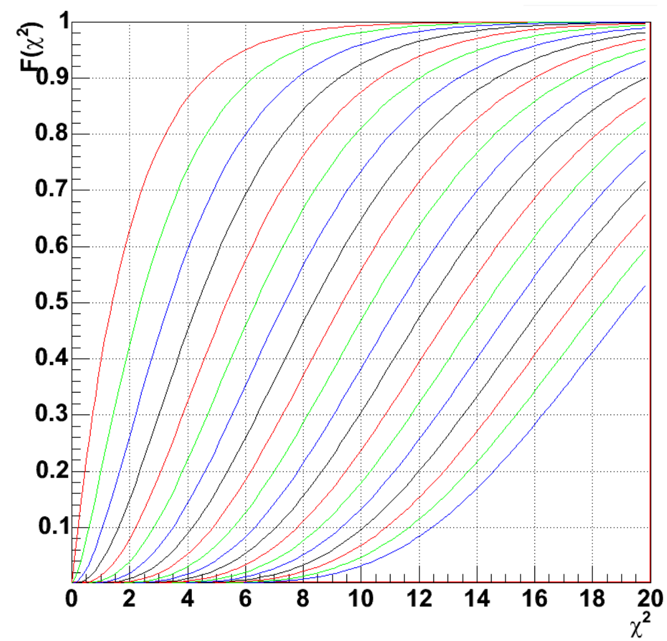
- **<u>Podsumowując</u>, w naszym przypadku musimy <u>dla danej realizacji próby (wyniku eksperymentu)</u> wyznaczyć wartość testową T i porównać ją z odpowiednim kwantylem rozkładu** χ^2 **o odpowiedniej liczbie stopni swobody:**

$$\boxed{T > \chi^2_{1-\alpha}}$$

- **Jeżeli ten warunek jest spełniony, to hipotezę odrzucamy** (punkty teoretyczne nie opisują danych eksperymentalnych na zadanym poziomie istotności)

- Skąd wziąć kwantyl? Z tablic lub z dystrybuanty:





Hipotezy zerowa i alternatywna

Błędy I i II rodzaju

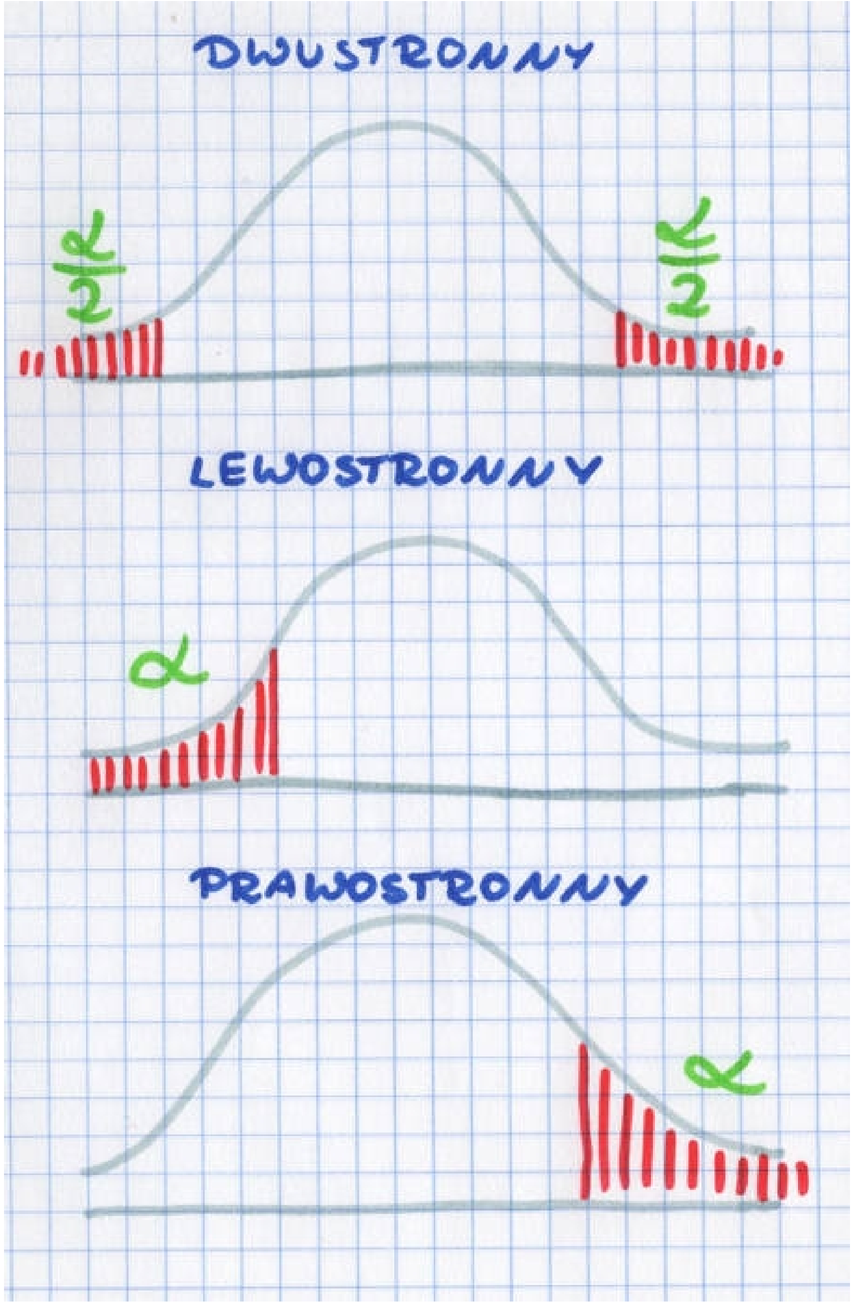
Na podstawie:
https://www.statystyczny.pl/hipotezy-statystyczne/

Hipoteza zerowa i alternatywna

- Hipoteza może być różna, przykłady hipotez:
  - średni wzrost Polaków to 175 cm
  - 2% dzieci w wieku szkolnym nie lubi czekolady
  - poziom szczęścia dwie minuty po zjedzeniu dużej porcji lodów jest wyższy niż przed zjedzeniem tejże dużej porcji lodów
- **Hipoteza zerowa** (H_0) to w uproszczeniu taka, gdy nie widzimy różnicy – np. czujemy się tak samo po zjedzeniu lodów jak przed
  - w przypadku pomiarów, np. wartość χ^2 jest mała → teoria opisuje dane
- **Hipoteza alternatywa** (H_A) to przeciwieństwo hipotezy zerowej, którą <u>możemy zdefiniować na kilka sposobów</u>, np.:
  - np. jest różnica w szczęściu w zjedzeniu lodów (test dwustronny)
  - poziom szczęścia po zjedzeniu lodów jest mniejszy (test jednostronny)
  - poziom szczęścia po zjedzeniu lodów jest większy (test jednostronny)
  - w przypadku pomiarów χ^2 jest duże → teoria nie opisuje danych (test jednostronny – rozkład χ^2 jest niesymetryczny)
- **Statystyka testowa** - to funkcja próby, na podstawie której wnioskuje się o odrzuceniu lub nie hipotezy statystycznej – wielkość mająca swój rozkład prawd.

Statystyka testowa

- **Statystyka testowa** - to funkcja próby, na podstawie której wnioskuje się o odrzuceniu lub nie hipotezy statystycznej – wielkość mająca swój rozkład prawdopodobieństwa
- Z naszej próby losowej (eksperymentu) dostajemy jedną wartość – ona znajduje się gdzieś w tym rozkładzie
- Obszar krytyczny (obszar odrzuceń) jest zawsze na końcu rozkładu
  - jeśli hipoteza mówi, że coś jest różne – dwustronny
  - mniejsze lub większe – jednostronny
- **Statystyka testowa ma swój (różny) rozkład zarówno dla H_0 jak i H_A!!!**



Błąd I rodzaju

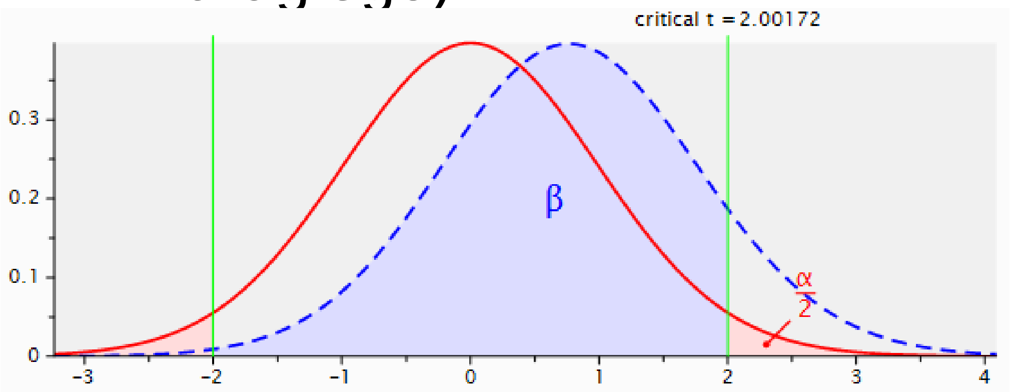
- **Błąd I rodzaju** to taki, gdzie odrzucamy hipotezę zerową a była ona prawdziwa
  - jeśli poziom szczęścia po zjedzeniu dużej porcji lodów zupełnie się nie zmienia i (<u>średnio w populacji</u>) jesteśmy tak samo szczęśliwi jak przed, a my odrzucimy tę hipotezę na rzecz hipotezy alternatywnej (<u>na podstawie próby losowej</u>), że poziom szczęścia się zmniejsza (np. przez te okropne wyrzuty sumienia, że znowu za dużo kalorii i będzie trzeba teraz dwie godziny ćwiczyć :)), to popełnimy właśnie błąd pierwszego rodzaju – bo odrzuciliśmy hipotezę zerową, która była prawdziwa
- Prawdopodobieństwo popełnienia błędu I rodzaju określa **poziom istotności** α, stąd oczekujemy, by był jak najmniejszy
- Oczywiście, na podstawie konkretnej próby losowej może się tak stać, że ten błąd popełnimy

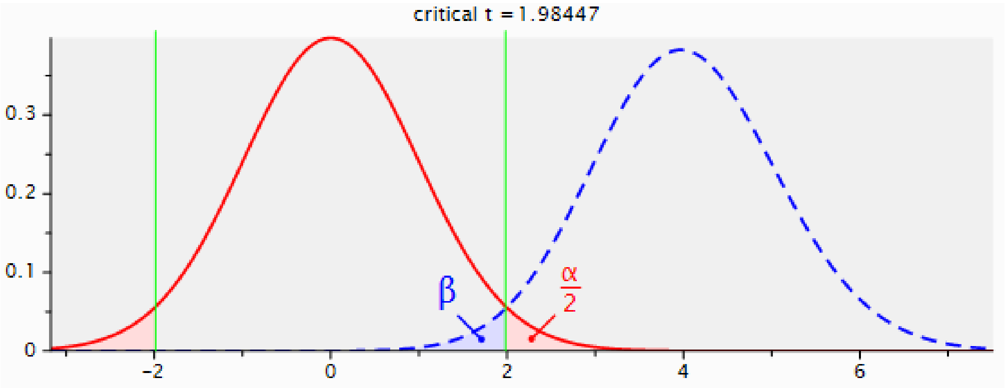
Błąd II rodzaju

- **Błąd II rodzaju** to taki, gdy nie odrzucimy hipotezy fałszywej
  - ma on miejsce w sytuacji, kiedy jednak ten poziom szczęścia przed zjedzeniem lodów i po zjedzeniu się różni. Jeśli (**średnio w społeczeństwie**) te wyrzuty sumienia z powodu zjedzenia dobrych i smacznych lodów powodują, że poziom szczęścia zdecydowanie spada po zjedzeniu lodów, a my stwierdzimy (**na podstawie próby losowej**), że nie ma żadnej różnicy, to wtedy popełniamy błąd II rodzaju. **Nie odrzucamy hipotezy zerowej, mimo że jest ona fałszywa.**
- **Prawdopodobieństwo** popełnienia błędu II rodzaju określamy jako β

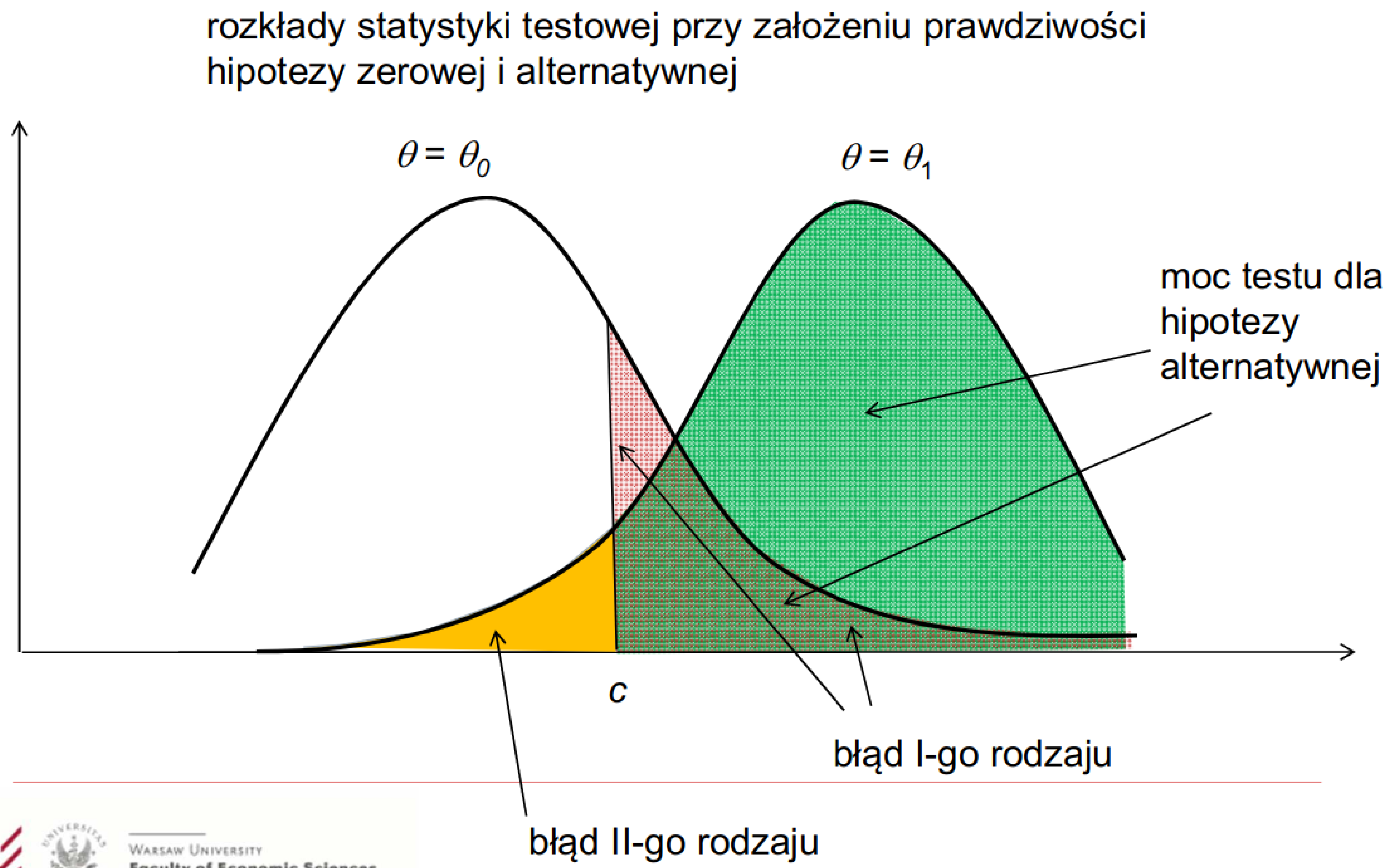
Moc testu

- Moc testu (prawdopodobieństwo, że prawidłowo odrzucimy hipotezę zerową) to 1-β. Inaczej mówiąc jest to prawdopodobieństwo niepopełnienia błędu II rodzaju.
- Moc testu zależy od kilku czynników:
  - Wielkości próby użytej w badaniu (im większa próba, tym większa moc testu).
  - Rzeczywistej wielkości efektu na tle losowej zmienności w populacji.
  - Przyjętego poziomu istotności α (między błędem I i II rodzaju jest taka zależność, że jeżeli zwiększamy prawdopodobieństwo popełnienia danego błędu, jednocześnie zmniejszamy je dla drugiego).





Moc testu:
interpretacja graficzna (1)



Przykład: winny czy inny

- **Przykład** działanie sądów może pomóc zrozumieć nam, kiedy przyjmujemy hipotezę alternatywną i dlaczego nieprzyjęcie hipotezy alternatywnej nie oznacza, że przyjmujemy hipotezę zerową.
- Wyobraźmy sobie Pana X oskarżonego o kradzież diamentu Pani Y. Pan X staje przed sądem.
  - **Hipotezą zerową** w tym przypadku jest niewinność Pana X.
    - Zakładamy, że Pan X wcale tych diamentów nie ukradł, że zrobiła to sprzątaczka, kucharka albo Pani Y schowała je w innej szufladzie i zupełnie o tym zapomniała.
  - **Hipoteza alternatywna**, to oczywiście wina Pana X. Skoro nie jest niewinny, to znaczy, że jednak ukradł diamenty i powinien zostać przez sąd skazany.
  - Założeniem jednak podstawowym jest brak winy – sąd musi znaleźć przekonujące argumenty, żeby móc Pana X oskarżyć.
    - **Jeśli je znajdzie** – to znaczy, że **odrzuca hipotezę zerową** (tą o niewinności) na rzecz hipotezy alternatywnej (że Pan X jest winny przestępstwa).
    - **Jeśli ich nie znajdzie**, to (nawet jeśli sąd wciąż będzie podejrzewać, że coś się w zachowaniu Pana X nie zgadza i że ta niewinność wcale nie jest zbyt pewna) **nie ma podstaw do odrzucenia hipotezy zerowej**. ***To wcale nie znaczy, że Pan X nie ukradł diamentów. To tylko oznacza, że sąd nie znalazł wystarczającego dowodu.***
- A gdzie tu błędy?
  - Jeśli Pan X ukradł diamenty i został skazany to sąd się nie pomylił. Jeśli diamentów nie ukradł i nie został uznany winnym, to również nie ma żadnego błędu.
  - Jeśli nie ukradł, a poszedł do więzienia, mamy do czynienia z błędem pierwszego rodzaju. Natomiast jeśli Pan X ukradł diamenty i został uniewinniony, to mamy do czynienia z błędem drugiego rodzaju.

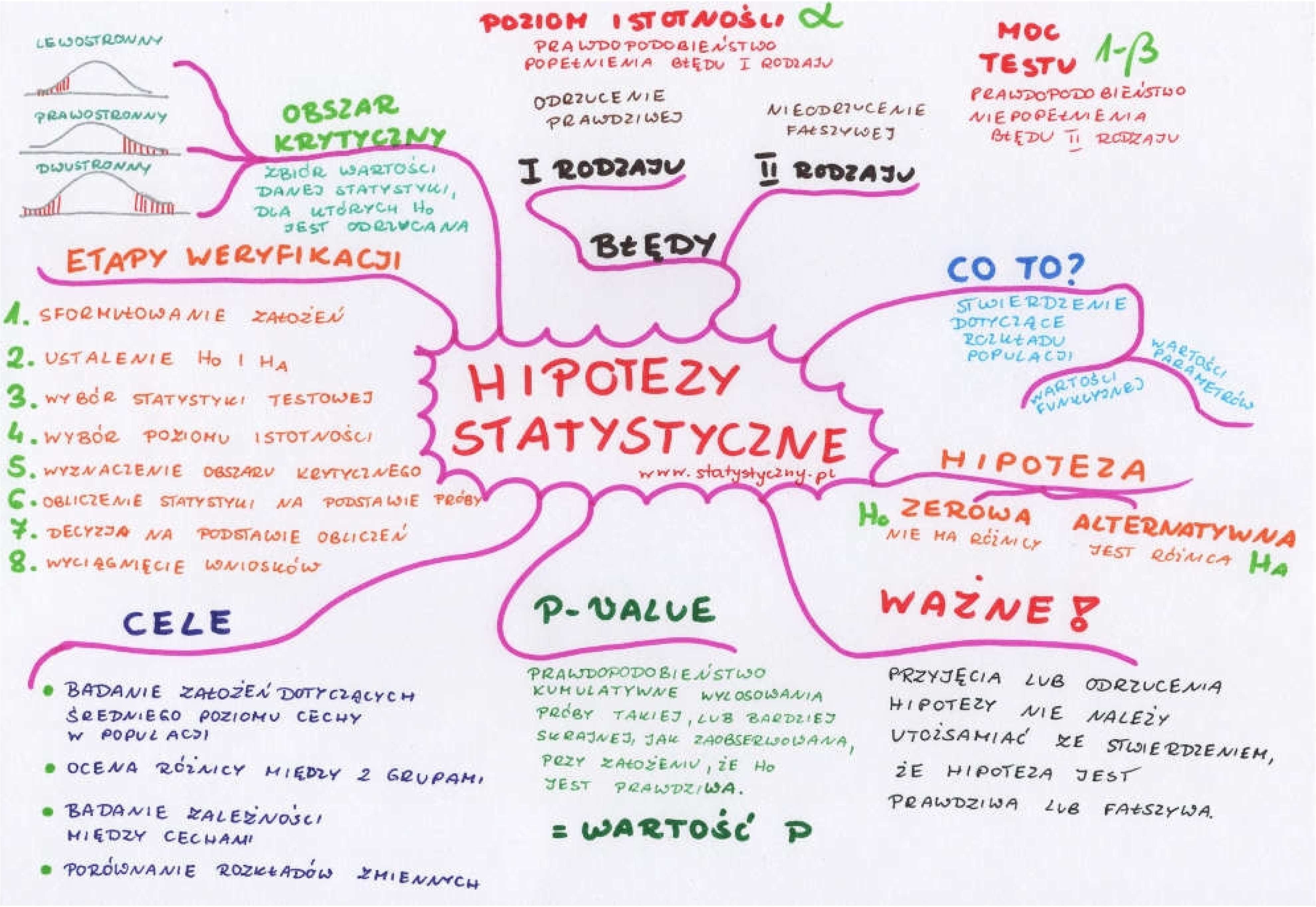
Błąd II rodzaju

- Ograniczenie błędów drugiego rodzaju jest bardzo istotne w niektórych testach. Np. w przypadku medycyny lepiej powiedzieć zdrowemu pacjentowi, żeby zrobił dodatkowe badanie (kiedy w przypadku zerowej hipotezy – pacjent jest zdrowy – wyszło nam błędnie podejrzenie choroby) niż chorego pacjenta odesłać do domu bez leczenia z informacją, że jest zdrowy (błąd drugiego rodzaju).

Weryfikacja hipotez

		DECYZJA	
		~~PRZYJĄĆ H_0~~ (NIE ODRZUCIĆ)	ODRZUCIĆ H_0
W RZECZYWISTOŚCI	H_0 PRAWDZIWA	DECYZJA PRAWIDŁOWA	BŁĄD I RODZAJU
	H_0 FAŁSZYWA	BŁĄD II RODZAJU	DECYZJA PRAWIDŁOWA

Metoda najmniejszych kwadratów



Dalsza lektura

- Polecam poniższe strony:
  - https://www.statystyczny.pl/hipotezy-statystyczne/
  - http://pogotowiestatystyczne.pl/istotnosc-statystyczna/
  - http://coin.wne.uw.edu.pl/azylicz/sm/sm09_2016.pdf

Metoda najmniejszych kwadratów
(least squares method)

Metoda najmniejszych kwadratów

- Jedną z najważniejszych metod estymacji parametrów jest zaproponowana przez Legendre'a i Gaussa
- Metoda ta jest szczególnym przypadkiem ogólniejszej metody największej wiarygodności (można ją z niej wyprowadzić)
- Założenia:
  - wynik pomiaru y_j przedstawiamy jako sumę nieznanej wielkości x oraz niepewności pomiarowej ε_j: $y_j = x + \epsilon_j$
  - dobieramy wielkości ε_j tak, aby ich suma kwadratów była najmniejsza: $\sum_j \epsilon_j^2 = \sum_j (x - y_j)^2 = min$
- Metoda ta może być również użyta, gdy wielkości y_j nie są wprost związane z x, lecz są to na przykład kombinacje liniowe (lub nieliniowe) wielu zmiennych $x_1, x_2, \ldots, x_n$
- **Warto przejrzeć rozdział o MNK w podręczniku Brandta, gdzie jest to wszystko bardzo dokładnie omówione**

Pomiary bezpośrednie

- Przedstawione wyżej założenia to najprostszy przypadek **pomiaru bezpośredniego o równej dokładności**

- Wykonujemy n pomiarów nieznanej wielkości x (np. długość stołu). Wyniki pomiarów obarczone są niepewnościami ε_j o których zakładamy, że opisane są rozkładem normalnym z wartością średnią równą zeru:

$$y_j = x + \epsilon_j \qquad E(\epsilon_j) = 0 \qquad E(\epsilon_j^2) = \sigma^2$$

- Zatem prawdopodobieństwo uzyskania wartości y_j jako wyniku pojedynczego pomiaru (wewnątrz małego przedziału $\mathrm{d}y$) wynosi:

$$f_j\, dy = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_j - x)^2}{2\sigma^2}\right) dy$$

- Logarytmiczna funkcja wiarygodności (dla n pomiarów):

$$l' = -\frac{1}{2\sigma^2} \sum_{j=1}^{n} (y_j - x)^2 + const$$

- Oczywiście, szukamy maksimum funkcji (warunek wiarygodności):

$$l = max$$

Pomiary bezpośrednie

- Możemy zauważyć, że warunek ten jest równoważny warunkowi (**najmniejszych kwadratów**):

$$M=\sum_{j=1}^{n}(y_j-x)^2=\sum_{j=1}^{n}\epsilon_j^2=min$$

- W tym przypadku estymatory:

$$\tilde{x}=\bar{y}=\frac{1}{n}\sum_{j=1}^{n}y_j \qquad \sigma^2(\bar{y})=\sigma^2/n$$

- W ogólniejszym przypadku, gdy mamy **różne dokładności** wyników pomiaru σ_j:

$$y_j=x+\epsilon_j \qquad E(\epsilon_j)=0 \qquad E(\epsilon_j^2)=\sigma_j^2=1/g_j$$

- **Wówczas warunek najmniejszych kwadratów wymaga dodatkowej wagi:**

$$M=\sum_{j=1}^{n}\frac{(y_j-x)^2}{\sigma_j^2}=\sum_{j=1}^{n}g_j(y_j-x)^2=\sum_{j=1}^{n}g_j\epsilon_j^2=min$$

- Wtedy estymatory:

$$\tilde{x}=\frac{\sum_{j=1}^{n}g_j y_j}{\sum_{j=1}^{n}g_j} \qquad \sigma^2(\tilde{x})=\left(\sum_{j=1}^{n}\frac{1}{\sigma_j^2}\right)^{-1}=\left(\sum_{j=1}^{n}g_j\right)^{-1} \qquad \boxed{\tilde{\epsilon}_j=y_j-\tilde{x}}$$

Średnia ważona z pomiarów o różnej dokł.

- Spodziewamy się, że wielkość $\tilde{\epsilon}_j = Y_j - \tilde{x}$ ma rozkład normalny z wartością średnią 0 i wariancją σ_j:
- Oczywiście wtedy wielkość $\tilde{\epsilon}_j/\sigma_j$ ma standardowy rozkład Gaussa
- Co to oznacza już dobrze wiemy, suma:
- Ma znany już nam rozkład χ^2 o n-1 stopniach swobody

$$M = \sum_{j=1}^{n} \left(\frac{\tilde{\epsilon}_j}{\sigma_j}\right)^2 = \sum_{j=1}^{n} \frac{Y_j - \tilde{x}}{\sigma_j^2} = \sum_{j=1}^{n} g_j (Y_j - \tilde{x})^2$$

- **Przykład:** średnia ważona z pomiarów o różnej dokładności
  - obliczamy wartość stałej ficzycznej (np. masy neutralnej cząstki K) poprzez średnią ważoną otrzymaną w różnych grupach eksperymentalnych

$$M = 7{,}2, \text{liczba st. swob. } n = 4 - 1 = 3, \alpha = 0.05, \chi^2_{0,95} = 7{,}82$$

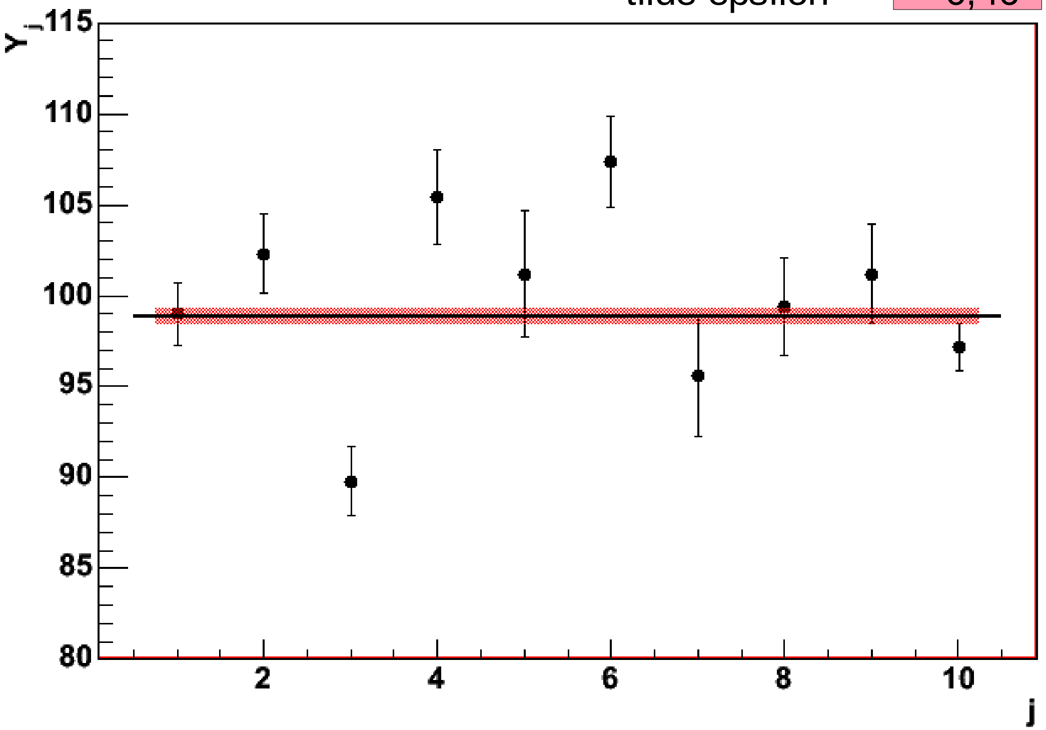
$$\tilde{x} = \sum_{j=1}^{4} \frac{Y_j g_j}{g_j} = 497{,}9$$

$$u(\tilde{x}) = \sqrt{\sigma^2(\tilde{x})} = = \left(\sum_{j=1}^{4} g_j\right)^{-1/2} = 0{,}2$$

j	y_j	σ_j	$g_j = 1/\sigma_j^2$	$y_j g_j$	$y_j - \tilde{x}$	$(y_j - \tilde{x})^2 g_j$
1	498.1	0.4	6.3	3038.0	0.2	0.3
2	497.4	0.33	10	4974.4	-0.46	2.1
3	498.9	0.5	4	1995.6	1.0	4.0
4	497.4	0.5	4	1989.8	-0.46	0.8
∑			24.3	11997.8		7.2

Średnia ważona z pomiarów o różnej dokł.

nr pomiaru	Yj	sigma_j	sigma_j^2	g_j	g_j x_j	Y_j – tilde x	e^2	e^2*g_j
1	99	1,7	2,89	0,35	34,26	0,19	0,04	0,01
2	102,3	2,2	4,84	0,21	21,14	3,49	12,19	2,52
3	89,8	1,9	3,61	0,28	24,88	-9,01	81,15	22,48
4	105,4	2,6	6,76	0,15	15,59	6,59	43,45	6,43
5	101,2	3,5	12,25	0,08	8,26	2,39	5,72	0,47
6	107,4	2,5	6,25	0,16	17,18	8,59	73,81	11,81
7	95,6	3,3	10,89	0,09	8,78	-3,21	10,29	0,95
8	99,4	2,7	7,29	0,14	13,64	0,59	0,35	0,05
9	101,2	2,7	7,29	0,14	13,88	2,39	5,72	0,78
10	97,2	1,3	1,69	0,59	57,51	-1,61	2,59	1,53
			sum(g_j)	2,18				
Ilość pomiarów	10		sum(x_j g_j)	215,12			M	47,02
			tilde x	98,81				
			tilde epsilon	0,46				



- Przeprowadzając test χ^2 na M widzimy, że hipotezę należy odrzucić:

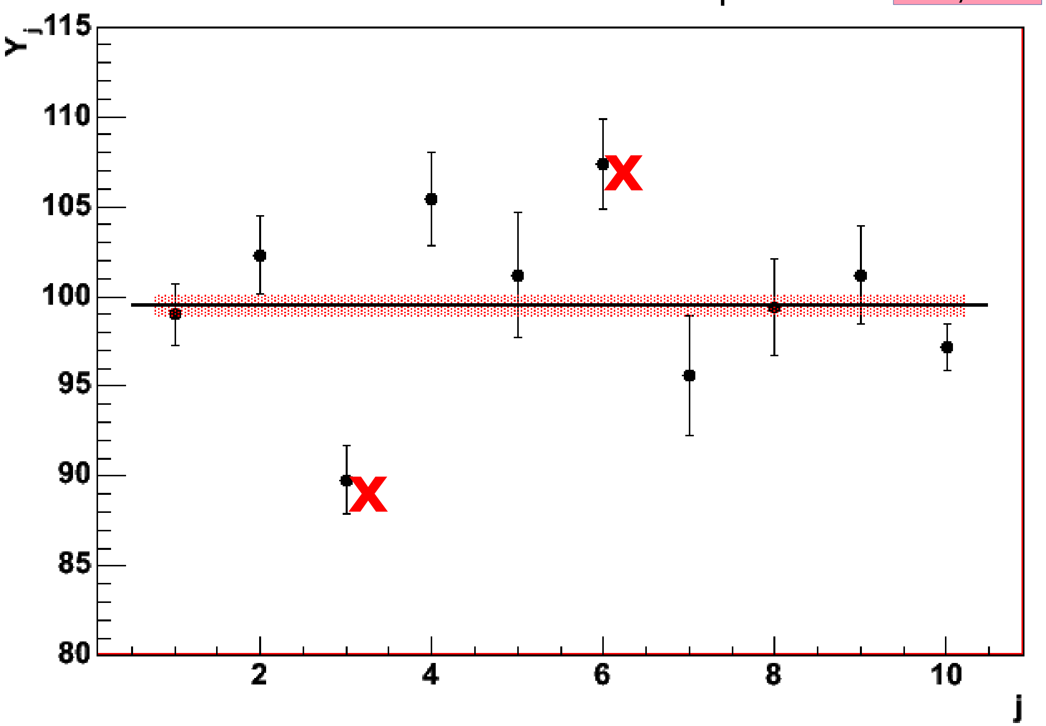
Kwantyle:

$$\chi^2_{0,9}(9) = 14{,}7$$
$$\chi^2_{0,95}(9) = 16{,}9$$
$$\chi^2_{0,99}(9) = 21{,}7$$

Przykład – odrzucenie pomiarów

nr pomiaru	Yj	sigma_j	sigma_j^2	g_j	g_j x_j	Y_j – tilde x	e^2	e^2*g_j
1	99	1,7	2,89	0,35	34,26	0,19	0,04	0,01
2	102,3	2,2	4,84	0,21	21,14	3,49	12,19	2,52
~~3~~								
4	105,4	2,6	6,76	0,15	15,59	6,59	43,45	6,43
5	101,2	3,5	12,25	0,08	8,26	2,39	5,72	0,47
~~6~~								
7	95,6	3,3	10,89	0,09	8,78	-3,21	10,29	0,95
8	99,4	2,7	7,29	0,14	13,64	0,59	0,35	0,05
9	101,2	2,7	7,29	0,14	13,88	2,39	5,72	0,78
10	97,2	1,3	1,69	0,59	57,51	-1,61	2,59	1,53

		sum(g_j)	1,74		
Ilość pomiarów	10	sum(x_j g_j)	173,06	M	12,73
		tilde x	99,45		
		tilde epsilon	0,57		



- Odrzucając pomiary najbardziej odbiegające od średniej mamy wynik spełniający test χ^2:

Kwantyle:

$$\chi^2_{0,9}(7) = 12,02$$
$$\chi^2_{0,95}(7) = 14,07$$
$$\chi^2_{0,99}(7) = 18,47$$

Pomiary pośrednie – przypadek liniowy

- Rozważymy teraz bardziej ogólny przypadek wielu (r) nieznanych wielkości x_i ($i=1,2,\ldots,r$) **mierzonych pośrednio**
- Interesujące nas wielkości fizyczne x nie podlegają pomiarom bezpośrednim, mierzymy natomiast liniowe kombinacje wielkości x_i mierzonych już bezpośrednio wielkości η_j:

$$\eta_j = p_{j0} + p_{j1} x_1 + p_{j2} x_2 + \ldots + p_{jr} x_r \qquad j=1,2,\ldots,n$$

wielkości mierzonych bezpośrednio

- Dla uproszczenia rachunków, można to zapisać inaczej:

$$f_j = \eta_j + a_{j0} + a_{j1} x_1 + a_{j2} x_2 + \ldots + a_{jr} x_r = 0$$

- W postaci wektorowej:

$$\boldsymbol{a_j} = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \ldots \\ a_{jr} \end{pmatrix} \qquad f_j = \eta_j + a_{j0} + \boldsymbol{a_j}^{\boldsymbol{T}} \boldsymbol{x} = 0$$

$$j=1,2,\ldots,n$$

- Jeśli wszystko zdefiniujemy wektorowo:

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_{j1} \\ \eta_{j2} \\ \ldots \\ \eta_{jr} \end{pmatrix} \qquad \boldsymbol{a}_0 = \begin{pmatrix} a_{10} \\ a_{20} \\ \ldots \\ a_{n0} \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nr} \end{pmatrix} \qquad \boldsymbol{f} = \boldsymbol{\eta} + \boldsymbol{a}_0 + A\boldsymbol{x} = 0$$

http://www.if.pw.edu.pl/~majanik/files/wiel.ps

Pomiary pośrednie – przypadek liniowy

- Rozważymy teraz bardziej ogólny przypadek wielu (r) nieznanych wielkości x_i ($i=1,2,\ldots,r$) **mierzonych pośrednio**

- Interesujące nas wielkości fizyczne x nie podlegają pomiarom bezpośrednim, mierzymy natomiast liniowe kombinacje wielkości x_i mierzonych już bezpośrednio wielkości η_j:

$$\eta_j = p_{j0} + p_{j1} x_1 + p_{j2} x_2 + \ldots + p_{jr} x_r \qquad j=1,2,\ldots,n$$

wielkości mierzonych bezpośrednio

- Dla uproszczenia rachunków, można to zapisać inaczej:

$$f_j = \eta_j + a_{j0} + a_{j1} x_1 + a_{j2} x_2 + \ldots + a_{jr} x_r = 0$$

- W postaci wektorowej:

$$\boldsymbol{a_j} = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \ldots \\ a_{jr} \end{pmatrix} \qquad f_j = \eta_j + a_{j0} + \boldsymbol{a_j}^{\boldsymbol{T}} \boldsymbol{x} = 0 \qquad j=1,2,\ldots,n$$

- Jeśli wszystko zdefiniujemy wektorowo:

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_{j1} \\ \eta_{j2} \\ \ldots \\ \eta_{jr} \end{pmatrix} \qquad \boldsymbol{a}_0 = \begin{pmatrix} a_{10} \\ a_{20} \\ \ldots \\ a_{n0} \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nr} \end{pmatrix} \qquad \boldsymbol{f} = \boldsymbol{\eta} + \boldsymbol{a}_0 + A\boldsymbol{x} = 0$$

Pomiary pośrednie – przypadek liniowy

- Oczywiście nadal zakładamy, że każdy pomiar obarczony jest niepewnością o rozkładzie normalnym:

$$y_j=\eta_j+\epsilon_j,\ \ E(\epsilon_j)=0,\ \ E(\epsilon_j^2)=\sigma_j^2=1/g_j \qquad\qquad \boldsymbol{y}=\boldsymbol{\eta}+\boldsymbol{\epsilon}$$

- Ponieważ zmienne y_j są zmiennymi niezależnymi, możemy wariancje przedstawić w postaci diagonalnej macierzy kowariancji:

$$C_y=C_\epsilon=\begin{pmatrix}\sigma_1^2 & 0 & \dots & 0\\ 0 & \sigma_2^2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \sigma_n^2\end{pmatrix} \qquad G_y=G_\epsilon=C_y^{-1}=C_\epsilon^{-1}=\begin{pmatrix}g_1 & 0 & \dots & 0\\ 0 & g_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & g_n\end{pmatrix}$$

- Wstawiając $\boldsymbol{y}=\boldsymbol{\eta}+\boldsymbol{\epsilon}$ do wzoru $\boldsymbol{f}=\boldsymbol{\eta}+\boldsymbol{a}_0+A\,\boldsymbol{x}=0$ otrzymujemy:

$$\boldsymbol{y}-\boldsymbol{\epsilon}+\boldsymbol{a}_0+A\,\boldsymbol{x}=0$$

- Rozwiązujemy ten układ ze względu na $\boldsymbol{x}$ stosując metodę największej wiarygodności (zakładając rozkład normalny pomiarów y_j). Wtedy:

$$M=\sum_{j=1}^{n}\frac{\epsilon_j^2}{\sigma_j^2}=\sum_{j=1}^{n}\frac{\left(y_j+\boldsymbol{a}_j^T\boldsymbol{x}+a_{j0}\right)}{\sigma_j^2}=\boldsymbol{\epsilon}^T G_y\,\boldsymbol{\epsilon}=\left(\boldsymbol{y}+\boldsymbol{a}_0+A\,\boldsymbol{x}\right)^T G_y\left(\boldsymbol{y}+\boldsymbol{a}_0+A\,\boldsymbol{x}\right)=min$$

Pomiary pośrednie – przypadek liniowy

- Jeśli wprowadzimy: $\boldsymbol{c}=\boldsymbol{y}+\boldsymbol{a}_0$
- Wówczas: $M=(\boldsymbol{c}+A\,\boldsymbol{x})^T G_y (\boldsymbol{c}+A\,\boldsymbol{x})=min$
- Można to dalej uprościć:

$$G_y=H^T H \qquad H=H^T=\begin{pmatrix} 1/\sigma_1 & 0 & \dots & 0 \\ 0 & 1/\sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/\sigma_n \end{pmatrix}$$

- Jeśli teraz wprowadzimy: $\boldsymbol{c}'=H\,\boldsymbol{c}$ $\quad A'=H\,A$
- Wówczas warunek nam się upraszcza: $M=(A'\,\boldsymbol{x}+\boldsymbol{c}')^2=min$
- Po rozwiązaniu dostajemy:

$$\tilde{\boldsymbol{x}}=-A^{+'}\boldsymbol{c}'$$

- W praktyce używamy wzoru: $\tilde{\boldsymbol{x}}=-(A'^T A')^{-1} A^T \boldsymbol{c}'=-(A^T G_y A) A^T G_y \boldsymbol{c}$
- Żeby wyznaczyć niepewności pomiarowe musimy policzyć macierz kowariancji: $G_{\tilde{x}}^{-1}=(A^T G_y A)^{-1}=(A'^T A')^{-1}$
- Pierwiaski kwadratowe z elementów diagonalnych to niepewności pomiarowe $\tilde{\boldsymbol{x}}$ (mimo, że $\boldsymbol{x}$ <u>nie podlegało bezpośredniemu pomiarowi</u>)

Pomiary pośrednie – przypadek liniowy

- Dla pomiarów bezpośrednich η_j:

$$\tilde{\boldsymbol{\epsilon}} = A\tilde{\boldsymbol{x}} + \boldsymbol{c} = -A(A^T G_y A)^{-1} A^T G_y \boldsymbol{c} + \boldsymbol{c}$$

$$\tilde{\boldsymbol{\eta}} = \boldsymbol{y} - \tilde{\epsilon} = \boldsymbol{y} + A(A^T G_y A)^{-1} A^T G_y \boldsymbol{c} - \boldsymbol{c}$$ - pomiary “poprawione”

$$\tilde{\boldsymbol{\eta}} = A(A^T G_y A)^{-1} A^T G_y \boldsymbol{c} - \boldsymbol{a}_0$$

$$G_{\tilde{\eta}}^{-1} = A(A^T G_y A)^{-1} A^T = A G_{\tilde{x}}^{-1} A^T$$

- Wzór $\eta_j = p_{j0} + p_{j1} x_1 + p_{j2} x_2 + \ldots + p_{jr} x_r$ będzie również prawdziwy dla estymatorów

Pomiary pośrednie – przypadek liniowy

- **Przykład:** dopasowanie prostej do zbioru pomiarów
  - mamy pomiary y_j zależne od pewnej zmiennej kontrolnej t_j (np. czasu)
  - zakładamy, że wartości zmiennej kontrolnej są dokładnie znane (zaniedbywane niepewności) – inaczej przypadek nieliniowy
  - zakładamy liniową postać: $\eta_j = y_j - \epsilon_j = x_1 + x_2 t_j$
  - i szukamy wielkości x mierzonych pośrednio
  - posługując się notacją macierzową: $\boldsymbol{\eta} - x_1 - x_2 \boldsymbol{t} = 0 \qquad \boldsymbol{a}_0 = 0$
  - czyli szukamy ostatecznie wektor: $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
  - wyniki pomiarów:

j	0	1	2	3
t_j	0.0	1.0	2.0	3.0
y_j	1.4	1.5	3.7	4.1
σ_j	0.5	0.2	1.0	0.5

Pomiary pośrednie – przypadek liniowy

- Obliczenia:

$$A = -\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \qquad \boldsymbol{y} = \boldsymbol{c} = -\begin{pmatrix} 1.4 \\ 1.5 \\ 3.7 \\ 4.1 \end{pmatrix}$$

$$G_y = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \qquad H = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$A' = -\begin{pmatrix} 2 & 0 \\ 5 & 5 \\ 1 & 2 \\ 2 & 6 \end{pmatrix} \qquad \boldsymbol{c}' = -\begin{pmatrix} 2.8 \\ 7.5 \\ 3.7 \\ 8.2 \end{pmatrix} \qquad A'^T \boldsymbol{c}' = -\begin{pmatrix} 62.2 \\ 94.1 \end{pmatrix}$$

Pomiary pośrednie – przypadek liniowy

- Obliczenia c.d.:

$$(A'^T A')^{-1} = -\begin{pmatrix} 34 & 39 \\ 39 & 65 \end{pmatrix}^{-1} = \frac{1}{689}\begin{pmatrix} 65 & -39 \\ -39 & 34 \end{pmatrix} = \begin{pmatrix} 0.0943 & -0.0556 \\ -0.0556 & 0.0493 \end{pmatrix}$$

$$\widetilde{\boldsymbol{x}} = -(A'^T A')^{-1} A^T \boldsymbol{c}' = -(A^T G_y A) A^T G_y \boldsymbol{c}$$

$$\widetilde{\boldsymbol{x}} = \begin{pmatrix} 0.0943 & -0.0556 \\ -0.0556 & 0.0493 \end{pmatrix}\begin{pmatrix} 63.2 \\ 94.1 \end{pmatrix} = \begin{pmatrix} 0.636 \\ 1.066 \end{pmatrix}$$

$$C_{\tilde{x}} = \begin{pmatrix} 0.0943 & -0.0556 \\ -0.0556 & 0.0493 \end{pmatrix} \qquad u(\widetilde{x}_1) = 0.307 \qquad u(\widetilde{x}_2) = 0.222$$

$$\tilde{\boldsymbol{\eta}} = -A\tilde{\boldsymbol{x}} = \begin{pmatrix} 0.636 \\ 1.702 \\ 2.768 \\ 3.834 \end{pmatrix}$$

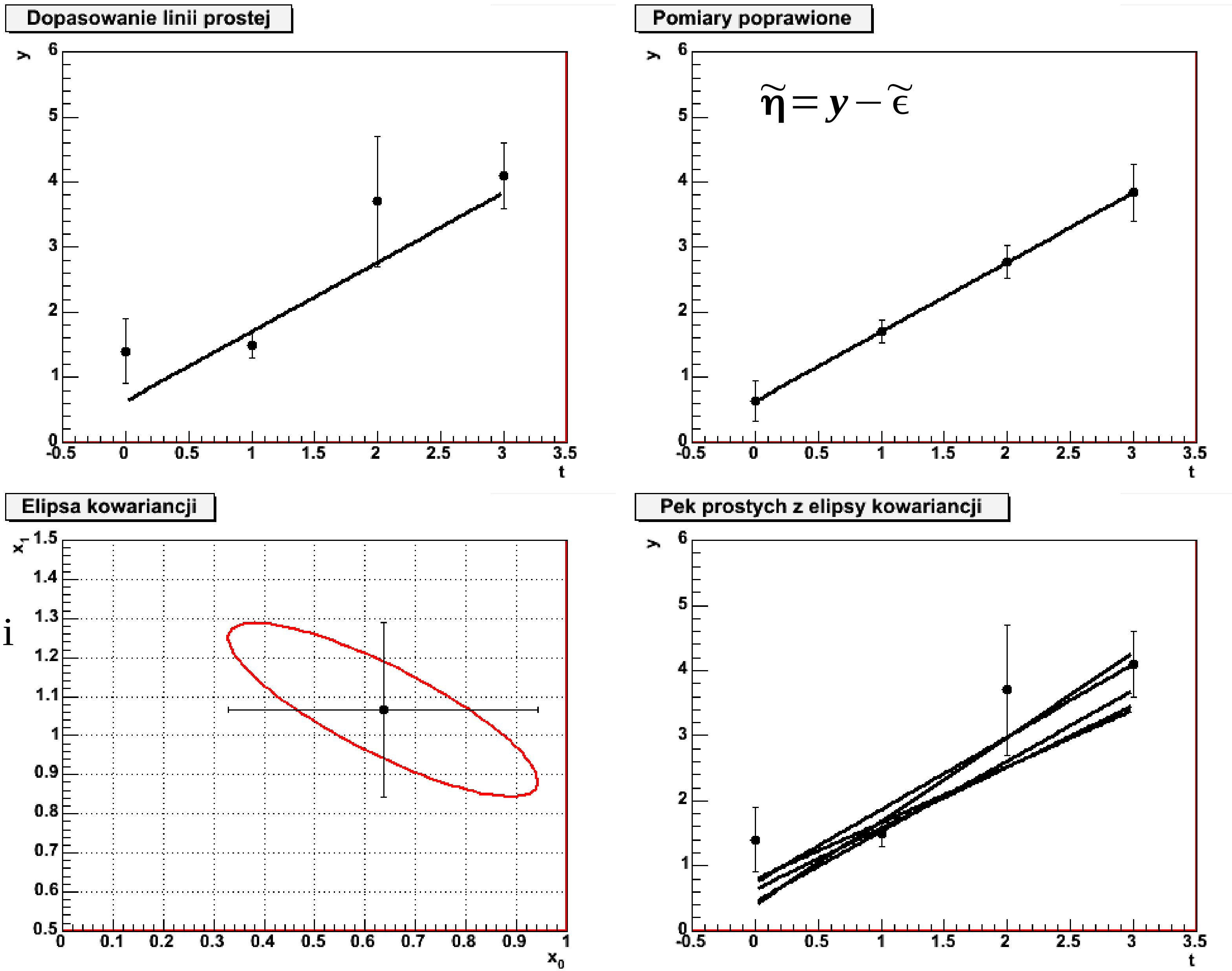
- Zminimalizowana suma kwadr.:

$$M = \left(\sum_{j=1}^{n} \frac{y_j - \widetilde{\eta}_j}{\sigma_j}\right)^2 = 4.507$$

N= 4 pomiary, 2 parametry, co daje
n-2 = r =2 stopnie swobody.
Zakładając poziom istotności 5% z tabel X^2:
$X^2_{0.95}$ = 5.99.
<u>Nie ma podstaw do odrzucenia hipotezy.</u>

Pomiary pośrednie – przypadek liniowy

- Wykresy:



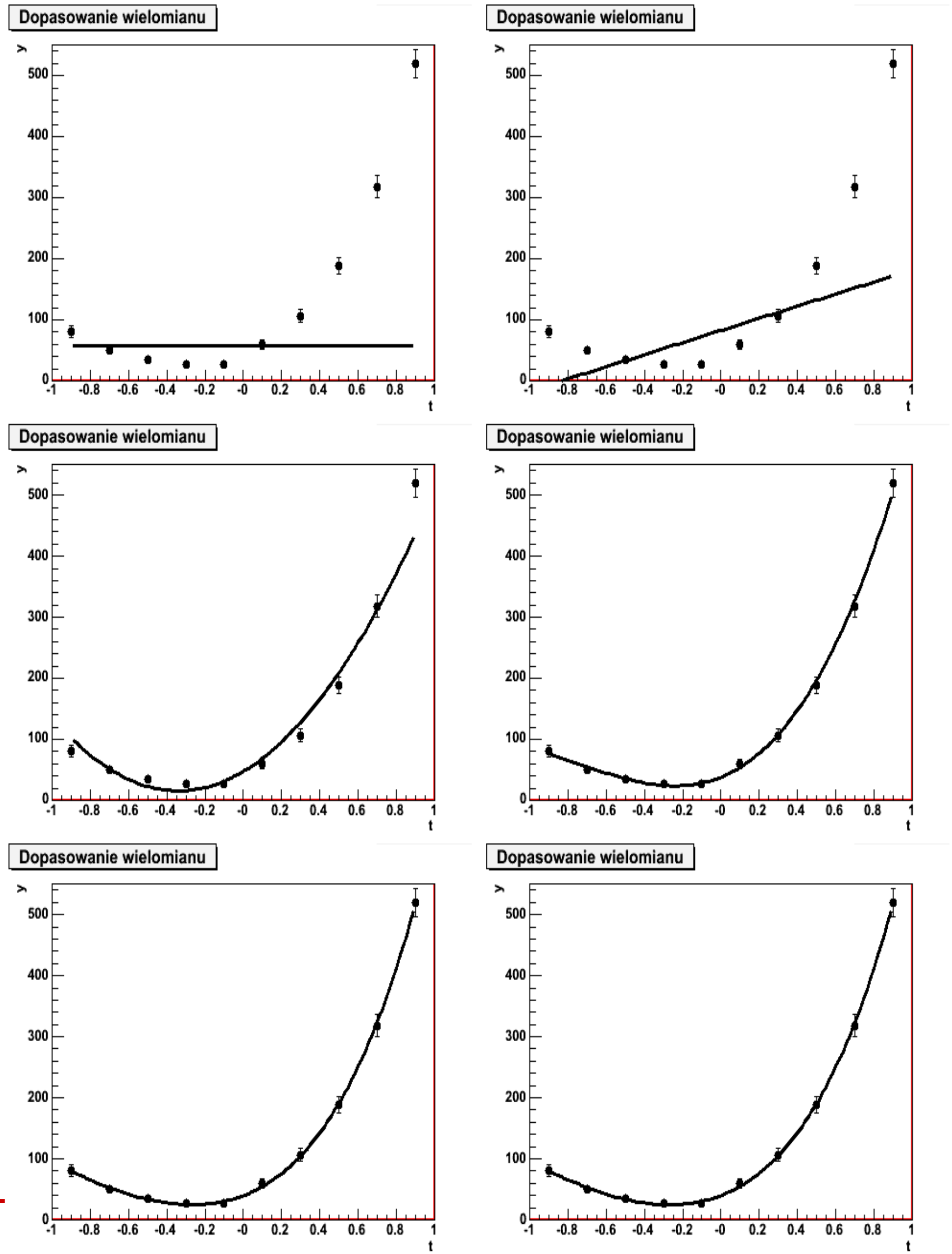
elipsa kowariancji
z macierzy kowariancji
$C_{\tilde{x}}$

Dopasowanie wielomianu

- Poprzedni przykład można uogólnić na wielomian wyższego rzędu:

$$\eta_j = h_j = x_1 + x_2 t_j + x_3 t_j^2 + \ldots + x_r t_j^{r-1}$$

$$\boldsymbol{a}_0 = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & t_1 & t_1^2 & \ldots & t_1^{r-1} \\ 1 & t_2 & t_2^2 & \ldots & t_2^{r-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & t_n & t_n^2 & \ldots & t_n^{r-1} \end{pmatrix}$$



- Pomiary:

j	1	2	3	4	5	6	7	8	9	10
t_j	-0,9	-0,7	-0,5	-0,3	-0,1	0,1	0,3	0,5	0,7	0,9
y_j	81	50	35	27	26	60	106	189	318	520

- Wynik:

r	~x_1	~x_2	~x_3	~x_4	~x_5	~x_6	f	M
1	57,85						9	833,55
2	82,66	99,1					8	585,45
3	47,27	185,96	273,61				7	36,41
4	37,95	126,55	312,02	137,59			6	2,85
5	39,62	119,1	276,49	151,91	52,6		5	1,69
6	39,88	121,38	273,19	136,57	56,9	16,73	4	1,66

KONIEC

Test t-Studenta

- Mamy zmienną losową X o rozkładzie normalnym. Pobieramy próbę losową o liczebności N i wartości średniej $\bar{X}$
- Wariancja wartości średniej: $\sigma^2(\bar{X}) = \sigma^2(X)/N$
- Dla dostatecznie dużych prób wartość średnia z próby (na mocy centralnego twierdzenia granicznego) ma rozkład normalny $(\hat{x}, \sigma(\bar{X}))$
- Zmienna $y = \dfrac{\bar{X} - \hat{x}}{\sigma(\bar{X})}$ ma standardowy rozkład normalny
- Na ogół nie znamy jednak odchylenia standardowego $\sigma^2(X)$
- Posługujemy się estymatorem wariancji:

$$s_X^2 = \frac{1}{N-1} \sum_{j=1}^{N} (X_j - \hat{x})^2 \qquad s_{\bar{X}}^2 = \frac{1}{N(N-1)} \sum_{j=1}^{N} (x_j - \bar{x})^2$$

- **Pytanie:** jak bardzo będziemy odbiegać od rozkładu Gaussa, jeżeli we wzorze na y zastąpimy odchylenie estymatorem?
- Dla uproszczenia, przyjmiemy, że $\hat{x} = 0$ (każdy rozkład Gaussa możemy przesunąć o wartość średnią)

Test t-Studenta

- Rozpatrzmy zmienną losową T zdefiniowaną następująco:

$$T = \bar{X} / s_{\bar{X}} = \bar{X} \cdot \sqrt{N} / S_X$$

- Wielkość $(N-1) s_X^2 = f s_X^2$ ma rozkład χ^2 o liczbie stopni swobody $f = N-1$
- Wzór na zmienną T zmieni się nam zatem następująco:

$$T = \bar{X} / s_{\bar{X}} = \bar{X} \cdot \sqrt{N} \cdot \sqrt{f} / \chi$$

- Dystrybuanta zmiennej T będzie określona wzorem:

$$F(t) = P(T < t) = P\left(\frac{\bar{X}\sqrt{N}\sqrt{f}}{\chi} < t\right) = \frac{\Gamma\left(\frac{1}{2}(f+1)\right)}{\Gamma\left(\frac{1}{2}f\right)\sqrt{\pi}\sqrt{f}} \int_{-\infty}^{t} \left(1 + \frac{t^2}{f}\right)^{\frac{1}{2}(f+1)} dt$$

- A odpowiadająca jej funkcja gęstości, nosząca nazwę **rozkładu t-Studenta**:

$$f(t) = \frac{\Gamma\left(\frac{1}{2}(f+1)\right)}{\Gamma\left(\frac{1}{2}f\right)\sqrt{\pi}\sqrt{f}} \left(1 + \frac{t^2}{f}\right)^{\frac{1}{2}(f+1)}$$

Rozkład t-Studenta

f=1
2
3
4
5
6
7
8
9
10

- Jak widać, rozkład t-Studenta jest symetryczny względem 0
- Rozkład dąży do rozkładu Gaussa gdy $f \to \infty$
- Z symetrii względem 0 mamy związek (analogicznie jak dla Gaussa): $P(|t| \leq t) = 2F(|t|) - 1$

- Możemy wyznaczyć graniczne wartości $\pm t'_{\alpha}$ odpowiadające poziomowi istotności α poprzez całkę:

$$\int_0^{t'_{\alpha}} f(t)\,dt = \frac{1}{2}(1-\alpha), \quad gdzie \quad t'_{\alpha} = t_{1-\frac{1}{2}\alpha}$$

- Kwantyle $t'_{\alpha} = t_{1-\frac{1}{2}\alpha}$ są stablicowane dla różnych poziomów istotności α oraz liczby stopni swobody *f*

- Jest to dwustronny test t-Studenta (jednostronny analogicznie)

Zastosowanie testu t-Studenta

- **Hipoteza:** zakładamy, że nasza populacja przewiduje wartość oczekiwaną z populacji mającej rozkład normalny równą λ_0
- Pobieramy próbę o liczebności N i wyznaczamy wartość średnią $\bar{X}$ oraz wariancję s_X^2
- Jeżeli przy założonym poziomie istotności α zachodzi nierówność:

$$|t| = \frac{|\bar{X} - \lambda_0|\sqrt{N}}{s_X} > t'_{\alpha} = t_{1-\frac{1}{2}\alpha}$$

- Wtedy **odrzucamy** naszą hipotezę
- W przypadku testu jednostronnego

$$t = \frac{(\bar{X} - \lambda_0)\sqrt{N}}{s_X} > t'_{2\alpha} = t_{1-\alpha}$$

Zastosowanie testu t-Studenta

- Powyższe rozważania możemy **uogólnić** na porównanie wartości średnich dwóch prób losowych z populacji X oraz Y o liczebnościach N_1 i N_2
- **Hipoteza:** równość wartości średnich z obu populacji: $\hat{x}=\hat{y}$
- Zakładamy (z centralnego twierdzenia granicznego), że wartości średnie z prób mają rozkład normalny z wariancjami:

$$\sigma^2(\bar{X})=\sigma^2(X)/N_{1,}\quad \sigma^2(\bar{Y})=\sigma^2(Y)/N_2$$

- Wariancje są estymowane przez estymatory:

$$s_{\bar{X}}^2=\frac{1}{N_1(N_1-1)}\sum_{j=1}^{N_1}(X-\bar{X})^2 \qquad s_{\bar{Y}}^2=\frac{1}{N_1(N_1-1)}\sum_{j=1}^{N_1}(Y-\bar{Y})^2$$

- Różnica wartości średnich z próby również ma rozkład zbliżony do normalnego: $\Delta=\bar{X}-\bar{Y}\Rightarrow\sigma^2(\Delta)=\sigma^2(\bar{X})+\sigma^2(\bar{Y})$
- Jeśli hipoteza jest prawdziwa, wówczas oczywiste jest, że $\hat{\Delta}=0$ oraz iloraz $\Delta/\sigma(\Delta)$ powinien podlegać rozkładowi Gaussa
- Tak postawiona hipoteza cicho zakłada, że X i Y to te same populacje

Test różnic t-Studenta

- Skoro tak, to oczywiście $\sigma^2(X)=\sigma^2(Y)$, zatem można je estymować za pomocą jednego estymatora jako średnią ważoną z dwóch prób:

$$s^2=\frac{(N_1-1)s_X^2+(N_2-1)s_Y^2}{(N_1-1)+(N_2-1)}$$

- Wtedy możemy zdefiniować estymatory:

$$s_{\bar{X}}^2=s^2/N_{1,}\quad s_{\bar{Y}}^2=s^2/N_{2,}\quad s_{\Delta}^2=s_{\bar{X}}^2+s_{\bar{Y}}^2=\frac{N_1+N_2}{N_1\cdot N_2}s^2$$

- Można udowodnić, że zmienna $\Delta/s(\Delta)$ podlega rozkładowi t-Studenta z liczbą stopni swobody $f=N_1+N_2-2$

- Równość wartości średnich można więc weryfikować posługując się **testem różnic Studenta**

- $\Delta/s(\Delta)$ obliczana jest na podstawie wyników dwóch prób. Jej wartość bezwzględną porównujemy z kwantylem rozkładu Studenta o liczbie stopni swobody *f* dla ustalonego poziomu istotności α. Sprawdzamy nierówność (**spełniona – odrzucamy hipotezę**):

$$|t|=\frac{|\Delta|}{s_\Delta}=\frac{|\bar{X}-\bar{Y}|}{S_\Delta}>t'_\alpha=t_{1-\frac{1}{2}\alpha}$$

Rozkład t-Studenta

f=1
2
3
4
5
6
7
8
9
10

- Jak widać, rozkład t-Studenta jest symetryczny względem 0
- Rozkład dąży do rozkładu Gaussa gdy $f \to \infty$
- Z symetrii względem 0 mamy związek (analogicznie jak dla Gaussa): $P(|t| \leq t) = 2F(|t|) - 1$

- Możemy wyznaczyć graniczne wartości $\pm t'_{\alpha}$ odpowiadające poziomowi istotności α poprzez całkę:

$$\int_0^{t'_\alpha} f(t)\,dt = \frac{1}{2}(1-\alpha), \quad gdzie \quad t'_\alpha = t_{1-\frac{1}{2}\alpha}$$

- Kwantyle $t'_\alpha = t_{1-\frac{1}{2}\alpha}$ są stablicowane dla różnych poziomów istotności α oraz liczby stopni swobody *f*
- Jest to dwustronny test t-Studenta (jednostronny analogicznie)

Test różnic t-Studenta - przykład

Numer pomiaru	Przyrząd 1	Przyrząd 2				
1	**100**	**97**	-0,21	0,04	-3,8	14,44
2	**101**	**101**	0,79	0,62	0,2	0,04
3	**102**	**102**	1,79	3,2	1,2	1,44
4	**100**	**99**	-0,21	0,04	-1,8	3,24
5	**98**	**101**	-2,21	4,89	0,2	0,04
6	**97**	**108**	-3,21	10,31	7,2	51,84
7	**100**	**101**	-0,21	0,04	0,2	0,04
8	**101**	**102**	0,79	0,62	1,2	1,44
9	**99**	**96**	-1,21	1,47	-4,8	23,04
10	**100**	**101**	-0,21	0,04	0,2	0,04
11	**98**		-2,21	4,89		
12	**101**		0,79	0,62		
13	**100**		-0,21	0,04		
14	**102**		1,79	3,2		
15	**103**		2,79	7,78		
16	**101**		0,79	0,62		
17	**99**		-1,21	1,47		
18	**100**		-0,21	0,04		
19	**102**		1,79	3,2		
Ilość pomiarów	19	10				
Średnia	**100,21**	**100,8**	**-0,59**			
Stopnie swobody	18	9				
S^2	43,16	95,6				
S^2/f	2,4	10,62				
S^2	49,1					
S^2 Delta	8,18					

- Mamy kwantyle:

$$t'_{0,2}(27)=t_{0,9}(27)=1.71$$

$$t'_{0,1}(27)=t_{0,95}(27)=2,05$$

$$t'_{0,02}(27)=t_{0,99}(27)=2,77$$

$$t'_{0,01}(27)=t_{0,995}(27)=3,05$$

$$t'_{0,004}(27)=t_{0,998}(27)=3,43$$

$$t'_{0,002}(27)=t_{0,999}(27)=3,69$$

- **Hipotezy nie można odrzucić**