

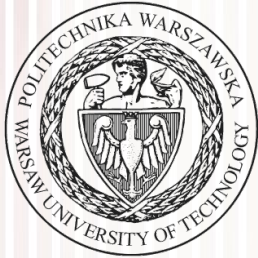


# Komputerowa analiza danych doświadczalnych

Wykład 11  
6.05.2016

dr inż. Łukasz Graczykowski  
[lgraczyk@if.pw.edu.pl](mailto:lgraczyk@if.pw.edu.pl)

*Semestr letni 2015/2016*



# Jednoczesna estymacja kilku parametrów - przykład

Weryfikacja hipotez  
statystycznych

Test F-Fischera

Test t-Studenta

Test  $\chi^2$



# Jednoczesna estymacja kilku parametrów - przykład

# Jednoczesna estymacja kilku parametrów

- **Przykład:** badamy zasięg cząstek  $\alpha$  w materii powstałych w wyniku rozpadu promieniotwórczego. Zasięg podlega rozkładowi normalnego wokół wartości średniej  $\rightarrow$  Innymi słowy, w wyniku eksperymentu mamy próbę losową o liczebności  $N$  z rozkładu normalnego

- Funkcja wiarygodności w tym przypadku:

$$L = \prod_{j=1}^N \frac{1}{\lambda_2 \sqrt{2\pi}} \exp \frac{-(X^{(j)} - \lambda_1)^2}{2\lambda_2^2}$$

$$l = \ln L = -\frac{1}{2} \sum_{j=1}^N \frac{(X^{(j)} - \lambda_1)^2}{\lambda_2^2} - N \ln \lambda_2 - \text{const}$$

- Układ równań wiarygodności:

$$\frac{\partial l}{\partial \lambda_1} = \sum_{j=1}^N \frac{X^{(j)} - \lambda_1}{\lambda_2^2} = 0$$

$$\frac{\partial l}{\partial \lambda_2} = \frac{1}{\lambda_2^3} \sum_{j=1}^N (X^{(j)} - \lambda_1)^2 - \frac{N}{\lambda_2} = 0$$

- Rozwiązanie:

$$\lambda_1 = \frac{1}{N} \sum_{j=1}^N X^{(j)}$$

$$\lambda_2 = \sqrt{\frac{\sum_{j=1}^N (X^{(j)} - \lambda_1)^2}{N}}$$

# Jednoczesna estymacja kilku parametrów

- Wyznaczymy macierz  $A$ , a następnie  $B$  i macierz  $C$  poprzez policzenie drugich pochodnych i w granicy  $N \rightarrow \infty$  zastąpienie ich wartościami oczekiwanymi:

$$\frac{\partial^2 l}{\partial \lambda_1^2} = -\frac{N}{\lambda_2^2} \quad \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_2} = -\frac{2 \sum X^{(j)} - \lambda_1}{\lambda_2^3} \quad \frac{\partial^2 l}{\partial \lambda_2^2} = -\frac{3 \sum (X^{(j)} - \lambda_1)^2}{\lambda_2^4} + \frac{N}{\lambda_2^2}$$

$$B = \begin{pmatrix} N/\lambda_2^2 & 0 \\ 0 & 2N/\lambda_2^2 \end{pmatrix} \quad C = B^{-1} = \begin{pmatrix} \lambda_2^2/N & 0 \\ 0 & \lambda_2^2/2N \end{pmatrix}$$

- Elementy diagonalne macierzy  $C$  to odchylenia standardowe estymowanych parametrów
- Elementy pozadiagonalne wynoszą 0, zatem brak jest korelacji między estymowanymi parametrami

$$-(l(\boldsymbol{\lambda}) - l(\tilde{\boldsymbol{\lambda}})) = 1/2 (\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}})^T A (\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}}) + \dots$$

$$L = k \exp(-1/2 (\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}})^T B (\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}})) \quad C = B^{-1}$$

$$B = E(A), \text{ dla } N \rightarrow \infty$$

$$-A = \begin{pmatrix} \frac{\partial^2 l}{\partial \lambda_1^2} & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_p} \\ \frac{\partial^2 l}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 l}{\partial \lambda_2^2} & \cdots & \frac{\partial^2 l}{\partial \lambda_2 \partial \lambda_p} \\ \frac{\partial^2 l}{\partial \lambda_p \partial \lambda_1} & \frac{\partial^2 l}{\partial \lambda_p \partial \lambda_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_p^2} \end{pmatrix}$$



# Weryfikacja hipotez statystycznych

# Weryfikacja hipotez statystycznych

- Podstawowe pojęcia:
  - **poziom istotności**  $\alpha$  – określa prawdopodobieństwo popełnienia błędu I rodzaju (błąd polegający na odrzuceniu hipotezy zerowej, która w rzeczywistości jest prawdziwa). Zwykle **maksymalne** prawdopodobieństwo (odrzucenia) ustalamy na 0,01, 0,03 lub 0,05
  - **poziom ufności**  $p=1-\alpha$  – określa prawdopodobieństwo wyznaczenia takiego przedziału ufności, że rzeczywista wartość parametru znajduje się w tym przedziale ufności (czyli jak bardzo ufamy naszemu wynikowi). Jest prawdopodobieństwem dopełniającym poziom istotności

# Weryfikacja hipotez statystycznych

- Co to jest i po co wykorzystujemy weryfikację hipotez statystycznych?
  - wyznaczamy **wartości parametrów**, które nie są całkowicie nieznane
  - mamy pewne **przypuszczenia (przewidywania, hipotezę)** co do ich wartości (np. z dotychczasowych badań czy modeli teoretycznych)
  - zadanie próby losowej (naszego pomiaru) to **weryfikacja (test)** tej hipotezy (przypuszczeń)
  - innymi słowy – procedura weryfikacji hipotez statystycznych to **test statystyczny**
- **Przykład:** pobrano 10-elementową próbę, uzyskaliśmy średnią arytmetyczną (jest ona zmienną losową)  $\bar{X} = 0,154$
- **Założenie (hipoteza):** zmienna losowa  $\bar{X}$  pochodzi z populacji opisanej rozkładem normalnym ze średnią 0 i odchyleniem  $1/\sqrt{10}$



# Weryfikacja hipotez statystycznych

- **Przykład:** pobrano 10-elementową próbę, uzyskaliśmy średnią arytmetyczną (jest ona zmienną losową)  $\bar{X}=0,154$
- **Założenie (hipoteza):** zmienna losowa  $\bar{X}$  pochodzi z populacji opisanej std. rozkładem normalnym ze średnią 0 i odchyleniem  $1/\sqrt{10}$
- Jakie jest prawdopodobieństwo zaobserwowania wartości  $|\bar{x}|\geq 0,154$  ?
- Można skorzystać z tabel dystrybuanty rozkładu normalnego oraz wprowadzonego kilka wykładów wcześniej wzoru:

$$P(|X - a| \leq n\sigma) = 2\Phi_0\left(\frac{nb}{b}\right) - 1 = 2\Phi_0(n) - 1$$

- Wtedy:  $P(|\bar{x}|\geq 0.154) = 2\{1 - \psi_0(0.154 \cdot \sqrt{10})\} = 0.62$
- Czyli, **jeżeli hipoteza jest prawdziwa**, istnieje prawdopodobieństwo 62%, że próba o liczebności 10 może mieć wartość średnią różniącą się więcej niż o 0,154 od wartości średniej 0 z populacji
- **Hipoteza jest prawdziwa czy nie?** Tak nie odpowiemy na to pytanie
- Źródłem trudności jest **probabilistyczny charakter wyników (próby)**

# Weryfikacja hipotez statystycznych

- Źródłem trudności jest **probabilistyczny charakter wyników (próby)**
- Postępujemy zatem inaczej:
  - przed przystąpieniem do analizy pobranej próby ustalamy określoną wartość prawdopodobieństwa  $\alpha$
  - następnie, zakładając, że hipoteza jest prawdziwa, pytamy czy prawdopodobieństwo zaobserwowania określonych wartości próby jest mniejsze niż  $\alpha$ ? W naszym przykładzie:  $P(|\bar{X}| \geq 0.154) < \alpha$  ?
  - **nierówność spełniona** – jest mało prawdopodobne, aby próba pochodziła z rozkładu określonego przez testowaną hipotezę → **możemy ją odrzucić**
  - **nierówność niespełniona** – omawiane prawdopodobieństwo przewyższa wartość  $\alpha$  → **nie ma podstaw do odrzucenia hipotezy** na zadanym poziomie ufności (to nie znaczy, że hipoteza na pewno jest prawdziwa! Hipoteza nie jest sprzeczna z wynikiem uzyskanym wskutek pobrania próby)

# Weryfikacja hipotez statystycznych

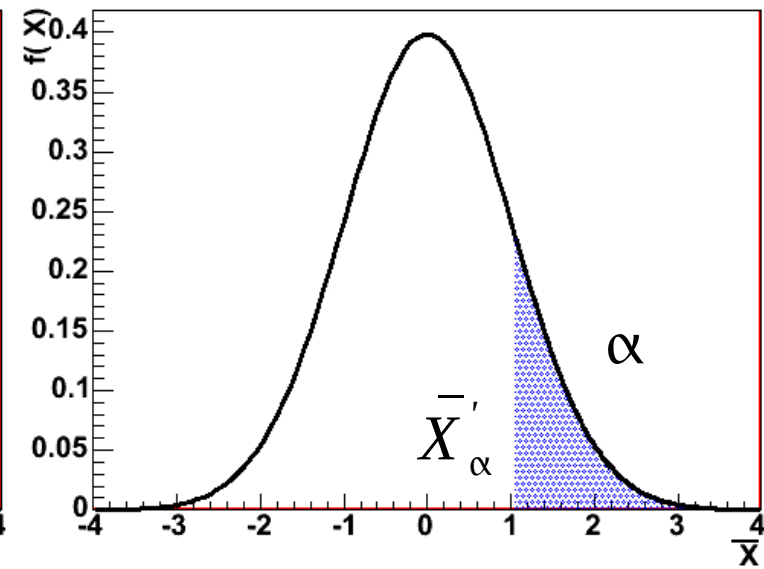
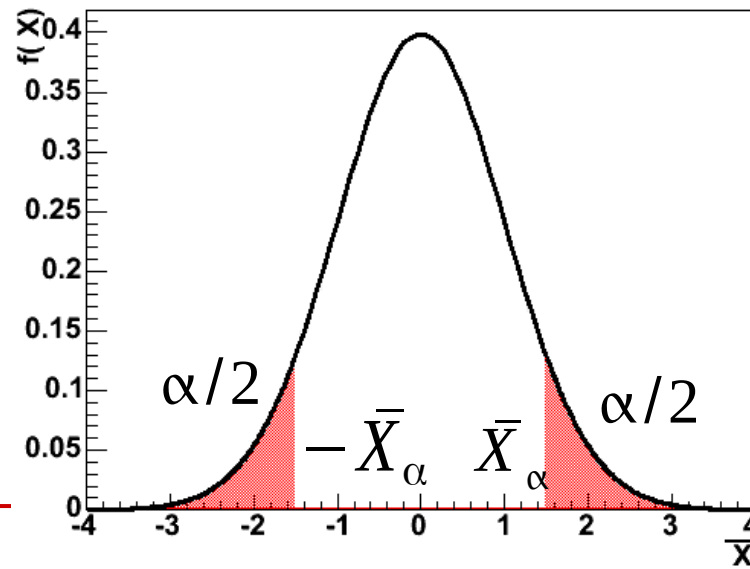
- Jak dobrać wartość prawdopodobieństwa (poziom istotności)  $\alpha$ ?
- Wszystko zależy od rozpatrywanego problemu (wybór jest subiektywny):
  - przy kontroli jakości ołówków wystarczy 1% (0,01)
  - w firmach ubezpieczeniowych może to być wartość za duża
  - w analizie danych empirycznych najczęściej przyjmuje się 5%, 1% lub 0,1%
- W naszym przykładzie na podstawie tablic dystrybuanty rozkładu normalnego można określić granice  $|\bar{X}|$  odpowiadające wyżej wymienionym poziomom istotności:
$$\alpha = 0,05 = 2 \{1 - \psi_0(1,96)\} = 2 \{1 - \psi_0(0,62 \cdot \sqrt{10})\}$$
$$\alpha = 0,01 = 2 \{1 - \psi_0(2,58)\} = 2 \{1 - \psi_0(0,82 \cdot \sqrt{10})\}$$
$$\alpha = 0,001 = 2 \{1 - \psi_0(3,29)\} = 2 \{1 - \psi_0(1,04 \cdot \sqrt{10})\}$$
- Do odrzucenia hipotezy wymagane jest, aby wartość  $|\bar{X}|$  przekraczała 0,62, 0,82, 1,04 dla powyższych poziomów istotności

# Weryfikacja hipotez statystycznych

- W niektórych przypadkach ważny jest również znak rozpatrywanej wielkości  $\bar{X}$
- W wielu procesach produkcji odstępstwa w jedną lub drugą stronę od założonych w projekcie parametrów wyrobu mogą powodować różne skutki. Na przykład piekarz produkujące zbyt duże bochenki zmniejsza swój zysk, natomiast produkując zbyt małe może narazić się na kary. Możemy zatem testować odchylenia w jednym kierunku i zapytać, czy spełniony jest warunek:

$$P(\bar{X} \geq x'_\alpha) < \alpha$$

- Jest to tzw. **test jednostrony**, w przeciwieństwie do **testu dwustronnego**



# Weryfikacja hipotez statystycznych

- W ogólnym przypadku weryfikację możemy przeprowadzać na podstawie innych wielkości niż wartość średnia
- Postępujemy najpierw w sposób następujący:
  - definiujemy funkcję próby najwygodniejszą dla prowadzenia testu, nazywamy ją statystyką  $T$
  - ustalamy poziom istotności  $\alpha$  i wyznaczamy taki podzbiór  $U$  z obszarami zmienności  $T$ , tak aby:  $P_H(T \in U) = \alpha$
  - wskaźnik  $H$  oznacza, że prawdopodobieństwo to obliczane jest przy założeniu, że hipoteza  $H$  jest prawdziwa
  - następnie pobieramy próbę i określamy wartość statystyki testowej  $T'$
  - jeżeli  $T'$  znajduje się wewnątrz obszaru krytycznego  $U$ , to hipotezę  $H$  można odrzucić
- Przedyskutujemy teraz kilka specyficznych testów



# Test równości wariancji (test $F$ -Fischera)

# Test równości wariancji ( $F$ -Fischera)

- **Częsty problem:** porównywanie wariancji populacji o jednakowych wartościach średnich
- **Przykład:** pomiar tej samej wielkości dwoma przyrządami pomiarowymi (zakładamy brak niepewności systematycznych – typu B)
- **Pytanie (hipoteza):** czy pomiary będą miały jednakowe wariancje (czy dokładność pomiaru jest jednakowa dla obu przyrządów)?
- Załóżmy, że rozważane populacje mają rozkład normalny
- Pobieramy próby o liczebności  $N_1$  i  $N_2$
- Dla każdej z pobranych prób wyznaczamy wariancję i liczymy iloraz
$$F = s_1^2 / s_2^2 \quad s^2(X) = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2 \quad s^2(\bar{X}) = \frac{1}{N \cdot (N-1)} \sum_{i=1}^N (X_i - \bar{X})^2 = s^2(X) / N$$
- Jeśli hipoteza o równości wariancji jest **prawdziwa**, to iloraz  $F$  powinien być bliski jedności  $F \sim 1$

# Test równości wariancji ( $F$ -Fischera)

- Jak pamiętamy, dla każdej próby losowej możemy również skonstruować statystykę o rozkładzie  $\chi^2$ :

$$\chi_1^2 = \frac{(N_1 - 1)s_1^2}{\sigma_1^2} = \frac{f_1 s_1^2}{\sigma_1^2} \qquad \chi_2^2 = \frac{(N_2 - 1)s_2^2}{\sigma_2^2} = \frac{f_2 s_2^2}{\sigma_2^2}$$

$f_1 = N_1 - 1$ ;  $f_2 = N_2 - 1$  - liczba stopni swobody

- Przy równych wariancjach iloraz  $F$ :  $F = \frac{f_2 \chi_1^2}{f_1 \chi_2^2}$
- Jak już wiemy, gęstość prawdopodobieństwa rozkładu  $\chi^2$  o liczbie stopni swobody  $f$  dana jest wzorem:

$$f(\chi^2) = k \cdot (\chi^2)^{\lambda-1} e^{-1/2\chi^2} \Rightarrow f(\chi^2) = \frac{1}{\Gamma\left(\frac{1}{2}f\right) 2^{\frac{1}{2}f}} (\chi^2)^{\frac{1}{2}(f-2)} e^{-\frac{1}{2}\chi^2}$$

- Możemy obliczyć prawdopodobieństwo  $W$  że wartość ilorazu  $\chi_1^2/\chi_2^2$  jest mniejsza niż pewna wartość  $Q$ :

$$W(Q) = P\left(\frac{\chi_1^2}{\chi_2^2} < Q\right) = \frac{\Gamma\left(\frac{1}{2}f\right)}{\Gamma\left(\frac{1}{2}f_1\right)\Gamma\left(\frac{1}{2}f_2\right)} \int_0^Q t^{\frac{1}{2}f_1-1} (t+1)^{-\frac{1}{2}f} dt \qquad f = f_1 + f_2$$



# Test równości wariancji ( $F$ -Fischera)

- Jeśli teraz wrócimy do ilorazu  $F$ , możemy go zapisać jako:

$$F = \frac{f_2}{f_1} Q$$

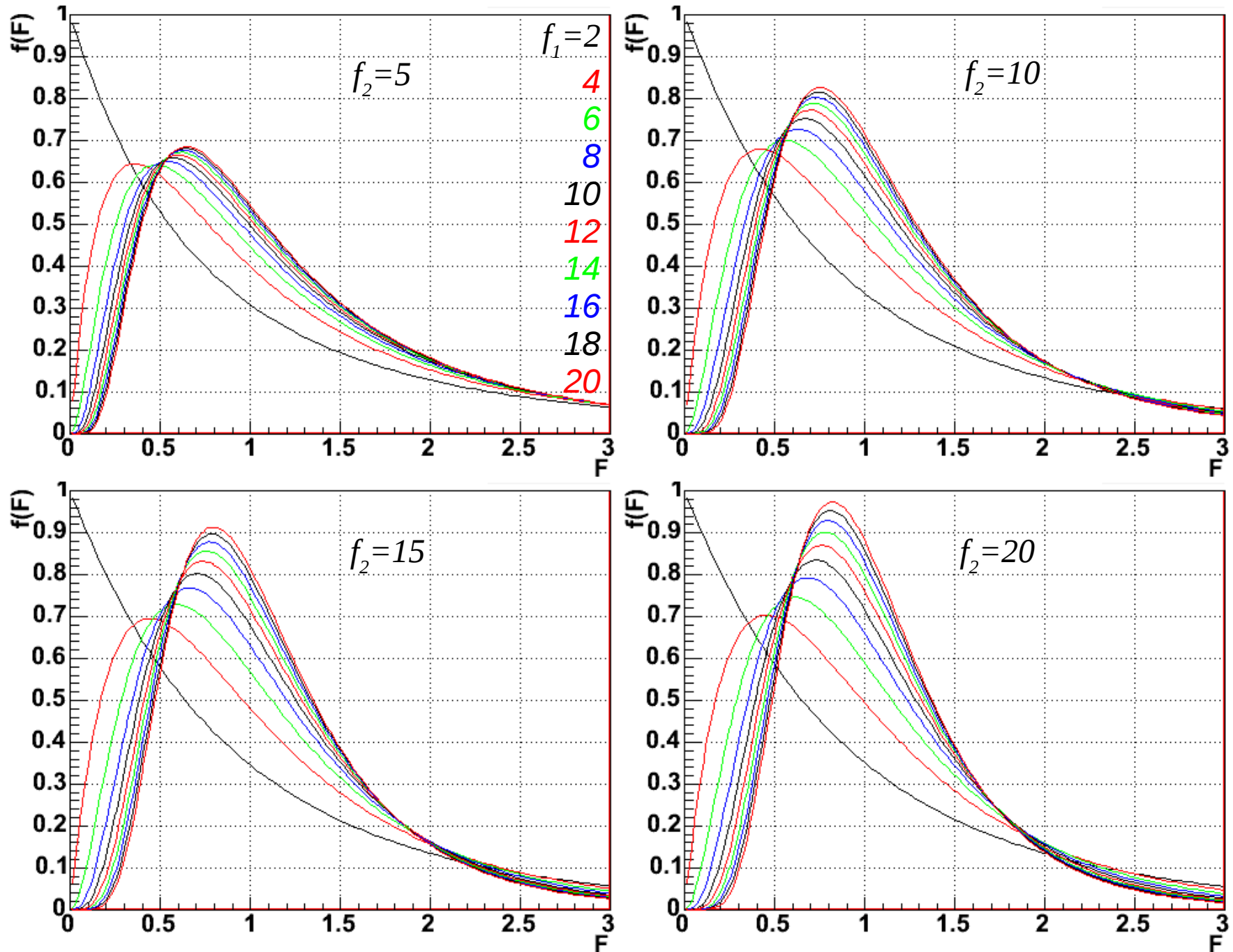
- Prawdopodobieństwo  $W$  zaś przyjmie postać:  $W(F) = P\left(\frac{S_1^2}{S_2^2} < F\right)$

- Ta forma jest tzw. **dystrybuantą rozkładu  $F$ -Fischera**

- **Rozkład  $F$ -Fischera** dany jest następującym wzorem:

$$f(F) = \left(\frac{f_1}{f_2}\right)^{\frac{1}{2}f_1} \frac{\Gamma\left(\frac{1}{2}(f_1+f_2)\right)}{\Gamma\left(\frac{1}{2}f_1\right)\Gamma\left(\frac{1}{2}f_2\right)} F^{\frac{1}{2}f_1-1} \left(1 + \frac{f_1}{f_2}F\right)^{-\frac{1}{2}(f_1+f_2)}$$

# Test równości wariancji ( $F$ -Fischera)



# Test równości wariancji ( $F$ -Fischera)

- Rozkład przypomina kształtem rozkład  $\chi^2$ , jest różny od 0 jedynie dla dodatnich wartości  $F$ , ponadto jest asymetryczny i szybko malejący z rosnącym  $F$

- Można udowodnić, że wartość oczekiwana dla  $f_2 > 2$ :  $E(F) = \frac{f_2}{f_2 - 2}$

- Szukamy teraz takiej granicznej wartości prawdopodobieństwa:

$$P\left(\frac{S_1^2}{S_2^2} > F'_\alpha\right) = \alpha$$

- Wielkość  $F'_\alpha$  to **kwantyl**  $F'_\alpha = F_{1-\alpha}$  rozkładu  $F$ :  $P\left(\frac{S_1^2}{S_2^2} < F'_\alpha\right) = P\left(\frac{S_1^2}{S_2^2} < F_{1-\alpha}\right) = 1 - \alpha$

- Jeżeli przy weryfikacji hipotezy uzyskamy wartość wariancji **wyższą** niż  $F'_\alpha = F_{1-\alpha}$ , to mówimy wówczas, że hipoteza  $\sigma_1^2 > \sigma_2^2$  jest **prawdziwa** z poziomem istotności (na poziomie istotności)  $\alpha$

- Wartości  $F'_\alpha = F_{1-\alpha}$  rozkładu  $F$ -Fischera są stabelaryzowane

- Pokazany wyżej test jest **testem jednostronnym**, możemy też zastosować **test dwustronny**

# Test równości wariancji ( $F$ -Fischera)

- W teście dwustronnym sprawdzamy 2 wartości graniczne:

$$P\left(\frac{S_1^2}{S_2^2} > F_{\alpha}''(f_1, f_2)\right) = \frac{1}{2} \alpha \qquad P\left(\frac{S_1^2}{S_2^2} < F_{\alpha}'''(f_1, f_2)\right) = P\left(\frac{S_2^2}{S_1^2} > F_{\alpha}''(f_2, f_1)\right) = \frac{1}{2} \alpha$$

- Relacja ta pozwala na posługiwanie się tabelaryzowanymi wartościami w obydwu przypadkach (dla testów jedno i dwustronnych)
- Dla wszystkich rozsądnych (zwykle używanych) wartości  $\alpha$  spełniony jest warunek:  $F_{1-\alpha/2} > 1$

- Czyli w praktyce musimy zweryfikować jedynie hipotezę:  $\frac{S_g^2}{S_k^2} > F_{1-\alpha/2}(f_g, f_k)$
- Indeksy  $g$  i  $k$  oznaczają większą i mniejszą wariancję z próby, czyli:

$$s_g^2 > s_k^2$$

- Jeżeli nierówność jest **spełniona**, to hipotezę o równości wariancji można **odrzuć**

# Test równości wariancji - przykład

Numer pomiaru	Przyrząd 1	Przyrząd 2				
1	100	97	0,2	0,04	-2,86	8,16
2	101	101	1,2	1,44	1,14	1,31
3	102	102	2,2	4,84	2,14	4,59
4	100	99	0,2	0,04	-0,86	0,73
5	98	101	-1,8	3,24	1,14	1,31
6	97	98	-2,8	7,84	-1,86	3,45
7	100	101	0,2	0,04	1,14	1,31
8	101		1,2	1,44		
9	99		-0,8	0,64		
10	100		0,2	0,04		
Średnia	99,8	99,86				
Stopnie swobody	9	6				
S <sup>2</sup>	19,6	20,86				
S <sup>2</sup> /f	2,18	3,48				
F	1,6					

- Korzystamy z kwantyli funkcji

$F$

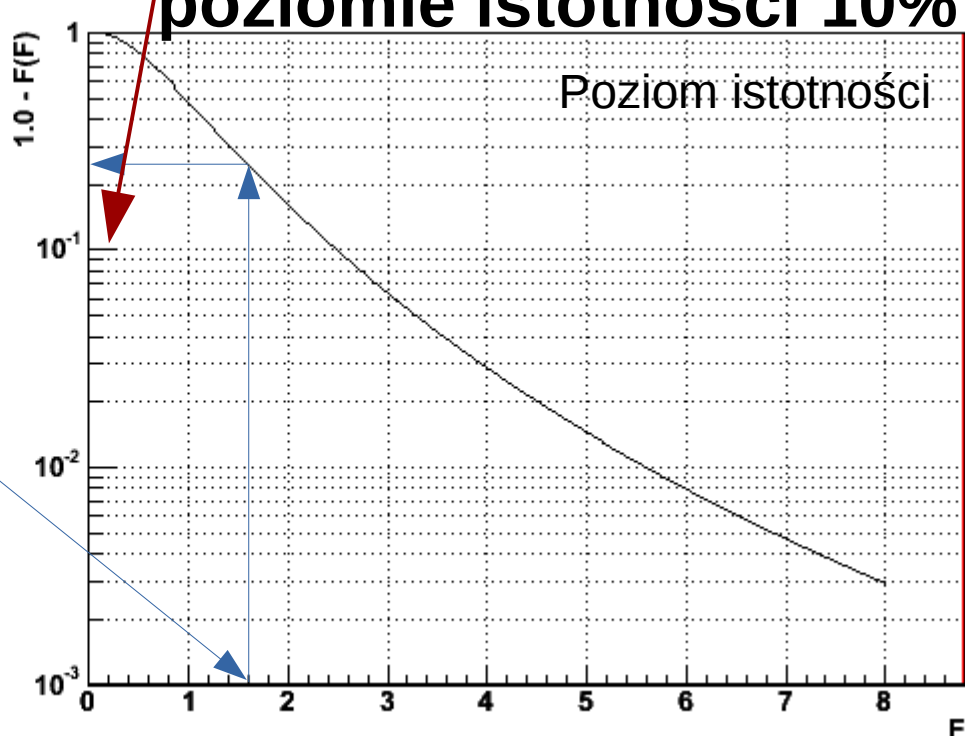
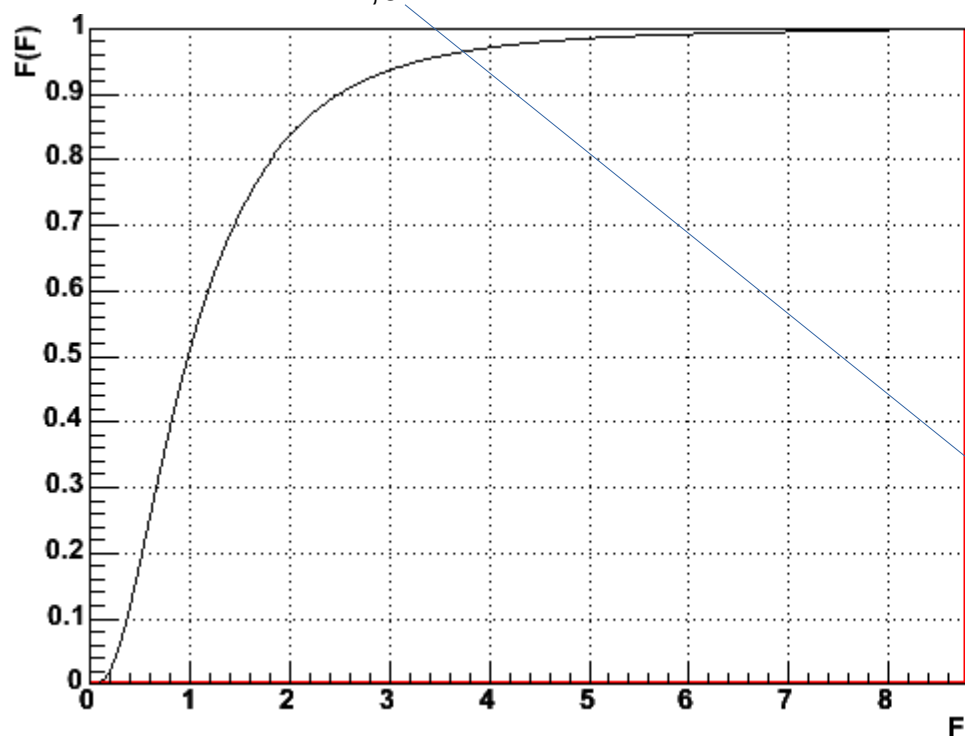
$$F''_{0,2}(6,9) = F_{0,9}(6,9) = 2.51$$

$$F''_{0,1}(6,9) = F_{0,95}(6,9) = 3.29$$

$$F''_{0,02}(6,9) = F_{0,99}(6,9) = 5.61$$

$$F''_{0,01}(6,9) = F_{0,995}(6,9) = 6.89$$

- 1.6 < 3.29 – nie ma podstaw do odrzucenia hipotezy na poziomie istotności 10%**





# Porównanie wartości średnich (test t-Studenta)

# 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  $y = \frac{\bar{X} - \hat{x}}{\sigma(\bar{X})}$  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  $(N-1)s_X^2 = fs_X^2$  ma rozkład  $\chi^2$  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:

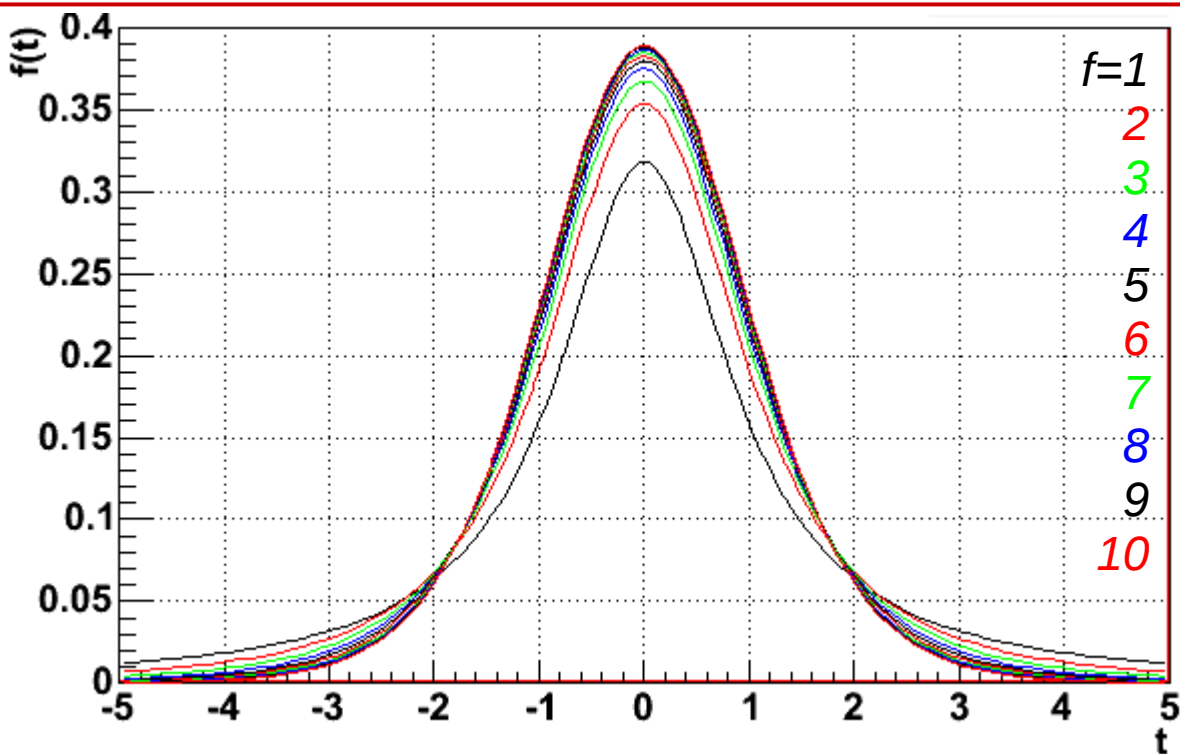
$$F(t) = P(T < t) = P\left(\frac{\bar{X} \sqrt{N} \sqrt{f}}{\chi} < t\right) = \frac{\Gamma\left(\frac{1}{2}(f+1)\right)}{\Gamma\left(\frac{1}{2}f\right) \sqrt{\pi} \sqrt{f}} \int_{-\infty}^t \left(1 + \frac{t^2}{f}\right)^{\frac{1}{2}(f+1)} dt$$

- 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



- Jak widać, rozkład t-Studenta jest symetryczny względem 0
- Rozkład dąży do rozkładu Gaussa gdy  $f \rightarrow \infty$
- 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  $\alpha$  poprzez całkę:

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

- Kwantyle  $t'_\alpha = t_{1 - \frac{1}{2}\alpha}$  są tablicowane dla różnych poziomów istotności  $\alpha$  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  $\lambda_0$
- Pobieramy próbę o liczebności  $N$  i wyznaczamy wartość średnią  $\bar{X}$  oraz wariancję  $s_X^2$
- Jeżeli przy założonym poziomie istotności  $\alpha$  zachodzi nierówność:

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

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

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

# 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:

$$s_{\bar{X}}^2 = \frac{1}{N_1(N_1-1)} \sum_{j=1}^{N_1} (X_j - \bar{X})^2 \quad s_{\bar{Y}}^2 = \frac{1}{N_2(N_2-1)} \sum_{j=1}^{N_2} (Y_j - \bar{Y})^2$$

- 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

- Tak postawiona hipoteza cicho zakłada, że  $X$  i  $Y$  to te same populacje
- 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  $\alpha$ . Sprawdzamy nierówność (**spełniona – odrzucamy hipotezę**):  $|t| = \frac{|\Delta|}{s_{\Delta}} = \frac{|\bar{X} - \bar{Y}|}{S_{\Delta}} > t'_{\alpha} = t_{1 - \frac{1}{2}\alpha}$

# 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
Stopnie swobody	18	9
S <sup>2</sup>	43,16	95,6
S <sup>2</sup> /f	2,4	10,62
S <sup>2</sup>	49,1	
S <sup>2</sup> Delta	8,18	

- Mamy kwantyle:

$$t'_{0,2}(27) = t_{0,9}(27) = 1,71$$

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

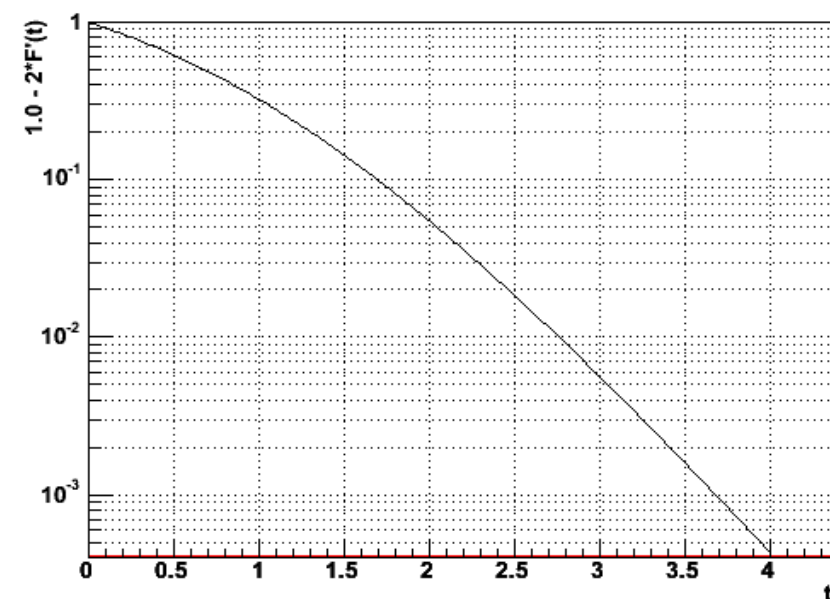
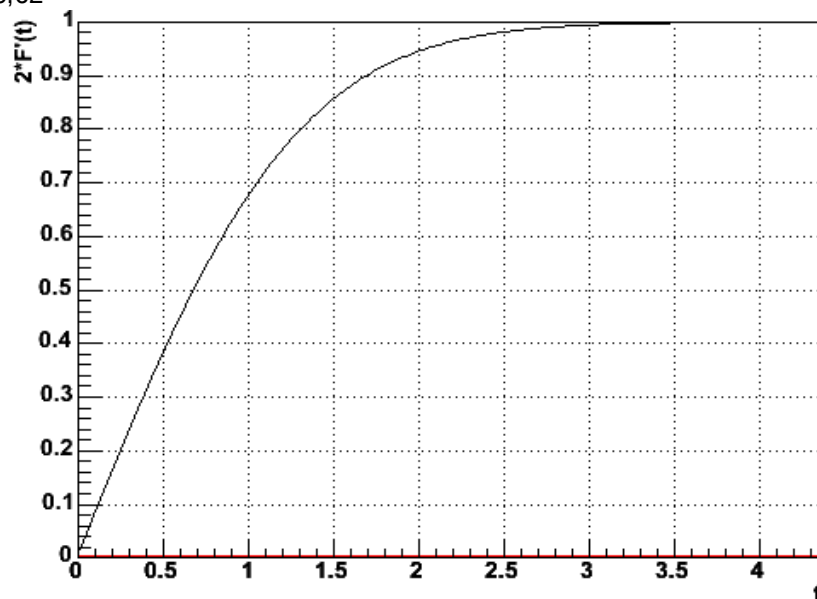
$$t'_{0,02}(27) = t_{0,99}(27) = 2,77$$

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

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

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

- Hipotezy nie można odrzucić





# Test dobroci $\chi^2$ dopasowania

# Test $\chi^2$ dobroci dopasowania

- Mamy  $N$  pomiarów  $g_i$ ,  $i=1, 2, \dots, N$  oraz ich niepewności  $\sigma_i$
- Wartość  $g_i$  jest wynikiem pomiaru wielkości prawdziwej (lecz nieznananej)  $h_i$ , spełniony jest zatem warunek:  $g_i = h_i + \epsilon_i$ ,  $i=1, 2, \dots, N$
- Wielkość  $\epsilon_i$  jest zmienną losową o rozkładzie normalnym o wartości średniej równej 0 i odchyleniu standardowym równym  $\sigma_i$
- **Weryfikujemy hipotezę taką, że:**  $h_i = f_i$ ,  $i=1, 2, \dots, N$  mają rozkład standardowy normalny  
 $u_i = \frac{g_i - f_i}{\sigma_i}$ ,  $i=1, 2, \dots, N$
- Tym samym rozkład  $\chi^2$  o  $N$  stopniach swobody będzie miała wielkość:  
$$T = \sum_{i=1}^N u_i^2 = \sum_{i=1}^N \left( \frac{g_i - f_i}{\sigma_i} \right)^2$$
- Gdy hipoteza jest fałszywa, wówczas poszczególne różnice wielkości mierzonych  $g_i$  oraz wielkości  $f_i$  (których dotyczy nasza hipoteza), dzielone przez niepewności  $\sigma_i$  przyjmują duże wartości
- Hipotezę **odrzucaamy**, gdy dla danego  $\alpha$  jest spełnione:  $T > \chi_{1-\alpha}^2$

# Liczba stopni swobody

- Często zarówno pomiary  $g_i$ , wartości prawdziwe  $h_i$ , jak i nasza hipoteza  $f_i$  wiążąca  $h_i$  z  $g_i$ , są funkcjami zmiennej kontrolnej  $t$ , której wartości znamy dokładnie:  $g_i = g_i(t_i)$ ,  $h_i = h_i(t_i)$ ,  $f_i = f_i(t_i)$
- Prosta hipoteza mówi, że jest liniowa zależność:  $f(t) = h(t) = at + b$
- Czyli mamy jakąś teorię, która mówi np. jak jakaś wielkość fizyczna zmienia się dajmy na to w czasie
- Często hipoteza (nasza teoria) określa wartości parametrów  $a$  i  $b$  i wtedy nasz problem ma  $N$  stopni swobody (wyniki pomiaru), jednak gdy parametry nie są znane, jedynym założeniem jest liniowość  $h(t)$
- W takim przypadku **każdy nieznany parametr, który estymujemy na podstawie pomiarów obniża liczbę stopni swobody o 1**
- Tworzymy więc estymatory parametrów  $a$  i  $b$  będące funkcjami pomiarów  $g_i$  oraz ich niepewności  $\sigma_i$ , wtedy nasza hipoteza przyjmuje postać:  $h_i = h(t_i) = f_i = \tilde{a} t_i + \tilde{b}$  i liczba stopni swobody wynosi  $N-2$  (w praktyce mamy 2 równania więzów między wielkościami  $u_i$ )



# Test $\chi^2$ i doświadczalny rozkład częstości

- Każdy ciągły rozkład prawdopodobieństwa  $f(x)$  można zamienić na dyskretny poprzez podział zakresu zmienności zmiennej losowej  $X$  na  $r$  przedziałów:  $\xi_1, \xi_2, \dots, \xi_i, \dots, \xi_r$

- Całkując  $f(x)$  w przedziałach otrzymujemy prawdopodobieństwo zaobserwowania zmiennej  $X$  w danym przedziale (binie)  $p_i$ :

$$p_i = P(x \in \xi_i) = \int_{\xi_i} f(x) dx; \quad \sum_{i=1}^r p_i = 1$$

- Z pobranej próby o liczebności  $n$  oznaczamy przez  $n_i$  elementy leżące w przedziale  $\xi_i$

- Oczywiście zachodzi relacja:  $n = \sum_{i=1}^r n_i$        $n_i = np_i$

- **Hipoteza:** zakładamy, że dla dużych wartości liczb  $n_i$  ich wariancja równa się  $n_i$  (patrz dyskusja o rozkładzie Poissona) i że rozkład wielkości  $u_i$ :

$$u_i^2 = \frac{(n_i - np_i)^2}{n_i}, \quad \text{lub} \quad u_i^2 = \frac{(n_i - np_i)^2}{np_i}$$

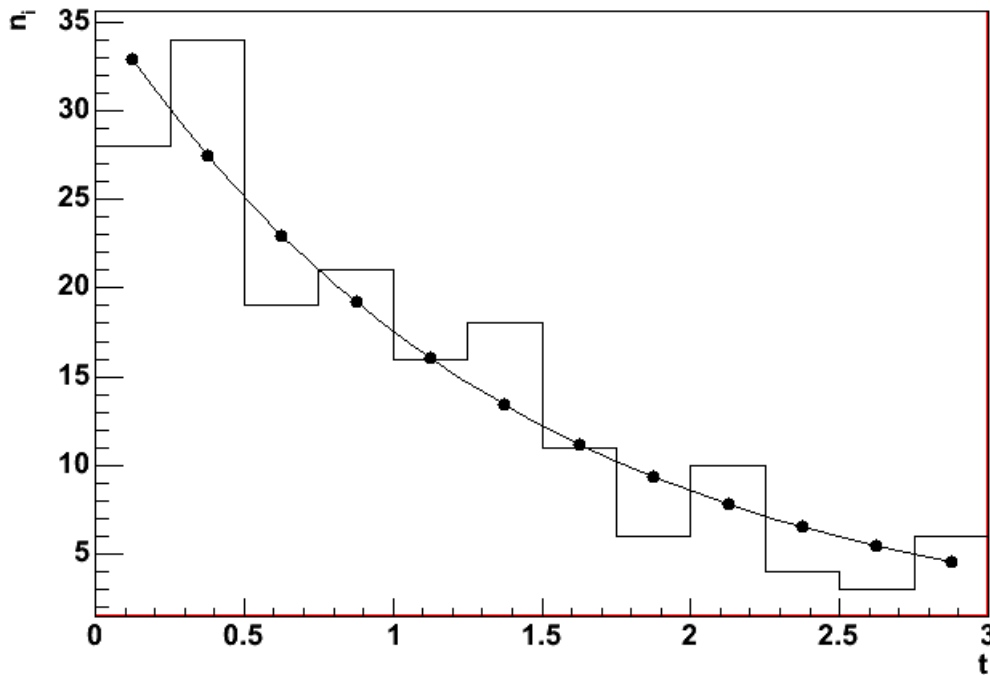
- dąży do rozkładu Gaussa

# Test $\chi^2$ i doświadczalny rozkład częstości

- Obliczamy teraz sumę kwadratów:  $\chi^2 = X^2 = \sum_{i=1}^r u_i^2$
- Na podstawie naszej hipotezy oczekujemy, że suma kwadratów (statystyka  $X^2$ ) ma asymptotycznie rozkład  $\chi^2$
- Ponieważ zmienne  $u_i$  są niezależne, więc liczba stopni swobody tego rozkładu równa się  $r-1$
- Oczywiście, jeżeli dodatkowo estymujemy  $p$  parametrów rozkładu na podstawie pomiarów, to liczba stopni swobody wynosi  $r-1-p$
- Wartość  $X^2$  porównujemy, tak jak do tej pory, z kwantylami rozkładu  $\chi^2$  dla danego poziomu istotności  $\alpha$ . Jeśli nierówność jest spełniona – odrzucamy hipotezę  $X^2 > \chi_{1-\alpha}^2$

# Test $\chi^2$ - przykład

zmierzone przypadki rozpadów



Numer binu	n_i	p_i	n p_i	(n_i - np_i)^2/np_i
1	28	0,1649	29,02	0,0360
2	34	0,1372	24,15	4,0202
3	19	0,1148	20,2	0,0718
4	21	0,0959	16,88	1,0065
5	16	0,0801	14,1	0,2567
6	18	0,0669	11,77	3,2917
7	11	0,0559	9,84	0,1371
8	6	0,0467	8,22	0,5992
9	10	0,0390	6,86	1,4328
10	4	0,0326	5,74	0,5262
11	3	0,0273	4,8	0,6779
12	6	0,0228	4,01	0,9841

n=sum(n\_i)    176     $\chi^2$     13,04

$$t_{0,9}(11) = 17,274$$

$$t_{0,95}(11) = 19,674$$

$$t_{0,99}(11) = 24,726$$

$$t_{0,995}(11) = 26,758$$

$$t_{0,998}(11) = 29,354$$

$$t_{0,999}(11) = 31,266$$

$$f(t) = \frac{1}{\tau} e^{-t/\tau}$$

$$\bar{t} = \int_0^{x_{max}} \frac{1}{\tau} e^{-t/\tau} = \tau$$

$$p_i = \int_{x_{min}}^{x_{max}} \frac{1}{\tau} e^{-t/\tau}$$

- Badamy stałą rozpadu promieniotwórczego.
- Estymatorem  $\tau$  jest wartość średnia  $t$
- Porównując kwantyle  $\chi^2$  widać, że nawet dla  $\alpha=0.1$  nie możemy odrzucić hipotezy



**KONIEC**