Komputerowa analiza danych doświadczalnych

Wykład 11
30.04.2019

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

Semestr letni 2018/2019

Weryfikacja hipotez statystycznych - ponownie

Test χ^2

Test F-Fishera

Test t-Studenta

Metoda najmniejszych kwadratów

Weryfikacja hipotez statystycznych

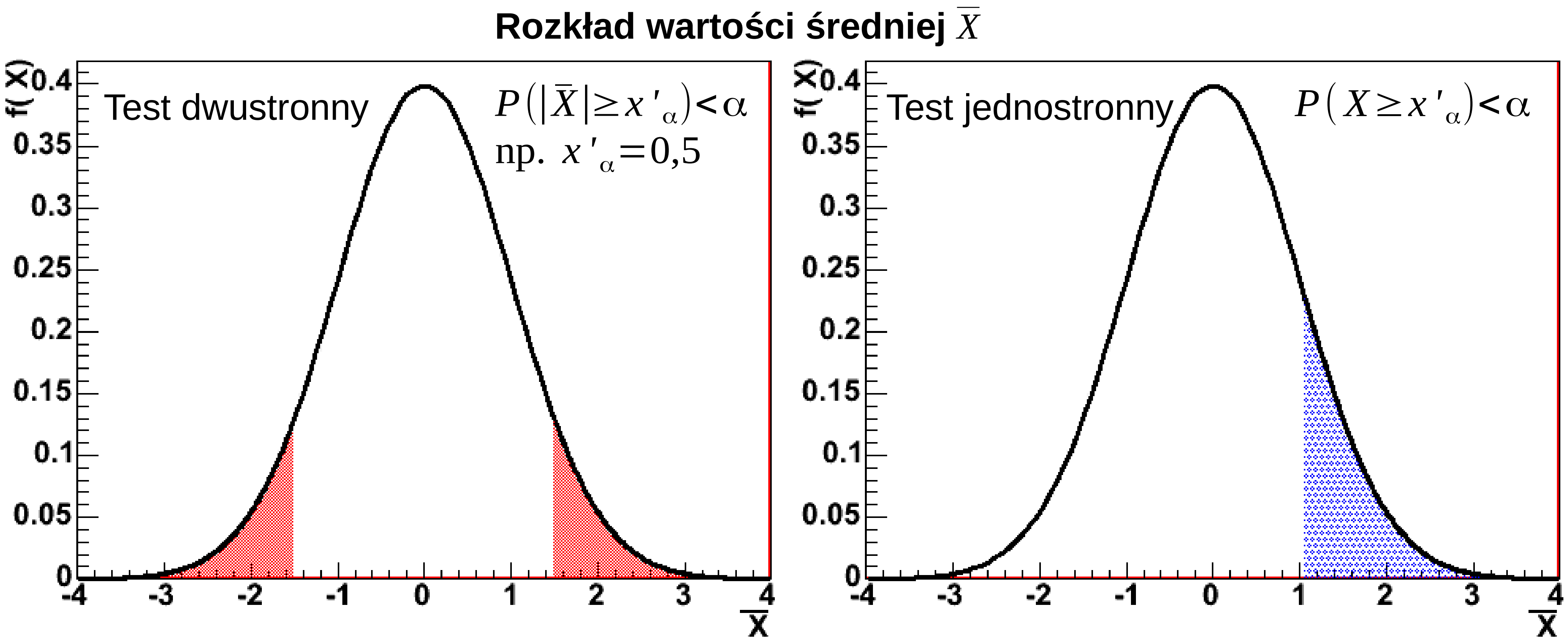
- **Przykład:** rozważamy zmienną losową X opisaną standardowym rozkładem Gaussa (średnia 0, odchylenie 1). Pobieramy 10-elementową próbę, uzyskaliśmy średnią arytmetyczną: $\bar{X} = 0{,}5$
- Jak na podstawie tej jednej realizacji próby (np. wyniku eksperymentu) możemy stwierdzić, czy pochodzi ona z 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 $1-\alpha$) 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



- 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 tej konkretnej próby **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: $\boxed{P(T > \chi^2_{1-\alpha}) = \alpha}$

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

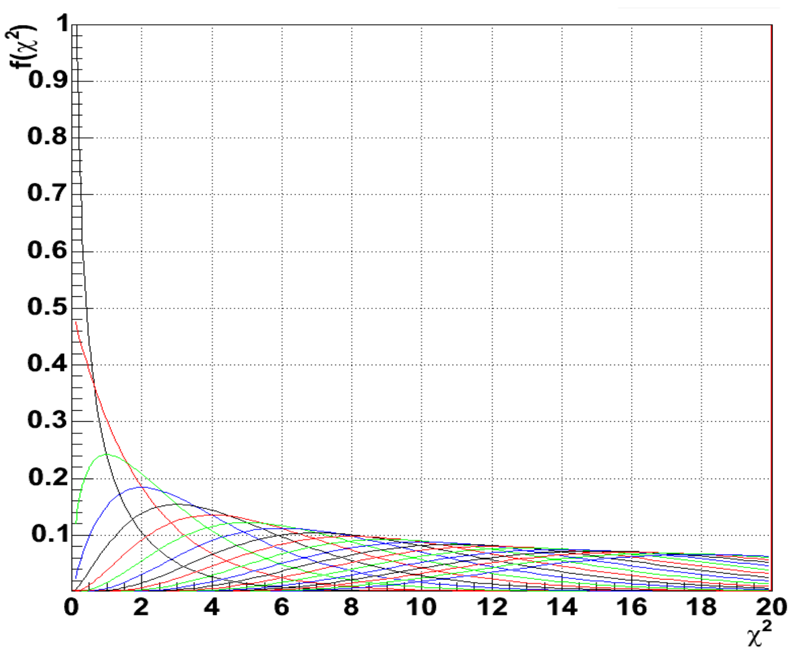
Test χ^2 dobroci dopasowania

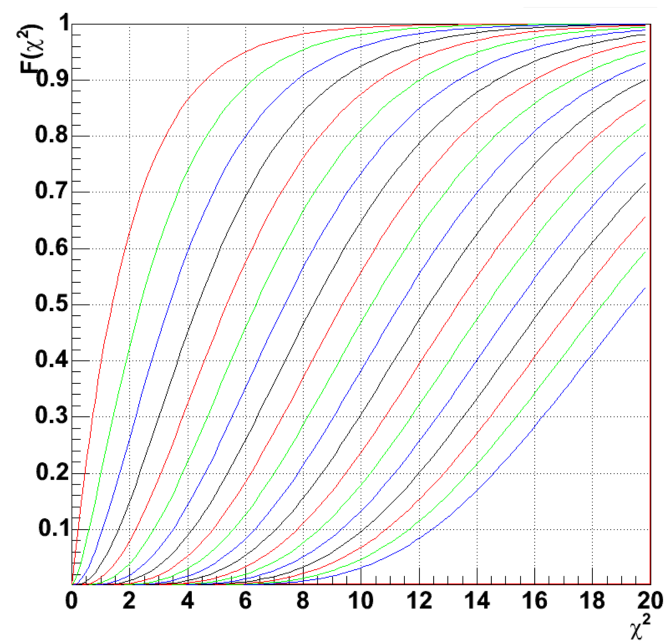
- **<u>Podsumowując</u>, w naszym przypadku musimy <u>dla danej realizacji próby (wyniku eksperymentu)</u> wyznaczyć wartość testową T i porównać ją z odpowiednim kwantylem rozkładu** χ^2 **o odpowiedniej liczbie stopni swobody:**

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

- **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:





Test χ^2 i doświadczalny rozkład częstości

- Możemy również rozważać zmienną losową X, (opisaną rozkładem $f(x)$) którą dzielimy na r przedziałów (histogram): $\xi_1, \xi_2, \ldots, \xi_i, \ldots, \xi_r$
- Całkując $f(x)$ w przedziałach otrzymujemy prawdopodobieństwo p_i zaobserwowania zmiennej X w danym przedziale (binie):
$$p_i = P(x \in \xi_i) = \int_{\xi_i} f(x)\,dx \; ; \quad \sum_{i=1}^{r} p_i = 1$$
- Z pobranej próby o liczebności n oznaczamy przez n_i elementy leżące w danym przedziale ξ_i
- Oczywiście zachodzi relacja: $n = \sum_{i=1}^{r} n_i$
suma wejść w poszczególnych binach równa jest liczebności próby
- **Oczekiwalibyśmy** (zakładając prawdziwość $f(x)$)**, że:** $n_i = np_i$
- **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: $u_i^2 = \frac{(n_i - n\,p_i)^2}{n_i}$, lub $u_i^2 = \frac{(n_i - n\,p_i)^2}{np_i}$ ma rozkład Gaussa

Test χ^2 i doświadczalny rozkład częstości

- Wtedy, suma kwadratów: $T=\sum_{i=1}^{r} u_i^2$
- Będzie miała (dla dużych n) rozkład χ^2
- **Jaka jest liczba stopni swobody?** Z defninicji histogramu mamy jedno **równanie więzów**: $n=\sum_{i=1}^{r} n_i$
- Zatem zmienne u_i **nie są niezależne**, więc liczba stopni swobody równa się r-1
- Oczywiście, jeżeli dodatkowo estymujemy p parametrów rozkładu na podstawie pomiarów (wprowadzamy p kolejnych wiezów uzależniających od siebie wielkości u_i), to liczba stopni swobody wynosi r-1-p
- Wartość T porównujemy, tak jak do tej pory, z kwantylami rozkładu χ^2 o określonej liczbie stopni swobody dla zadanego poziomu istotności α: $T>\chi^2_{1-\alpha}$
- **Jeśli nierówność jest spełniona – odrzucamy hipotezę**

Test χ^2 - przykład

Zadanie

Weryfikacja hipotez statystycznych (5 pkt.)

- Przeprowadzono eksperyment naświetlania wodorowej komory pęcherzykowej wiązką fotonów w celu badania oddziaływań fotonów z protonami. Fotony powodują powstawanie par elektron-pozyton, które mogą być wykorzystane do monitorowania wiązki fotonów. Częstość występowania zdjęć z 0,1,2,... parami elektron-pozyton powinna podlegać rozkładowi Poissona. Należy wczytać dane z pliku (w pierwszej kolumnie znajduje się liczba par elektronowych na zdjęciu k, a w drugiej liczba zdjęć zawierających k par elektronowych). Widzimy, że rozkład ten przypomina rozkład Poissona - próbujemy zatem obliczyć estymator największej wiarygodności dla parametry rozkładu Poissona (patrz Wykład 10 slajd 13) (1 pkt.)
- Narysować na jednym wykresie punkty pomiarowe i dopasowanie (metodą estymatora największej wiarygodności).
- Sprawdzić jakość dopasowania za pomocą testu χ2. W tym celu należy zaimplementować funkcję obliczającą statystykę testową χ2 zgodnie z wzorem $T = \sum_k \frac{(n_k - np_k)^2}{np_k}$

gdzie: nk - liczba obserwacji w k-tym binie, npk - przewidywana przez teorię liczba przypadków w k-tym binie

- Określić liczbę stopni swobody i obliczyć wartość statystyki testowej. (1 pkt.)
- Zaimplementować funkcję zwracającą wynik testu χ2 na zadanym poziomie istotności α

Wykorzystując zaimplementowaną funkcję zweryfikować hipotezę mówiacą, że dane pomiarowe podlegają rozkładowi Poissona. Dobrać odpowiednią wartość poziomu istotności. Uwaga! Kwanyl możemy odczytać z policzonej na ostatnich zajęciach dystrybuanty. (2 pkt.)

Zadanie

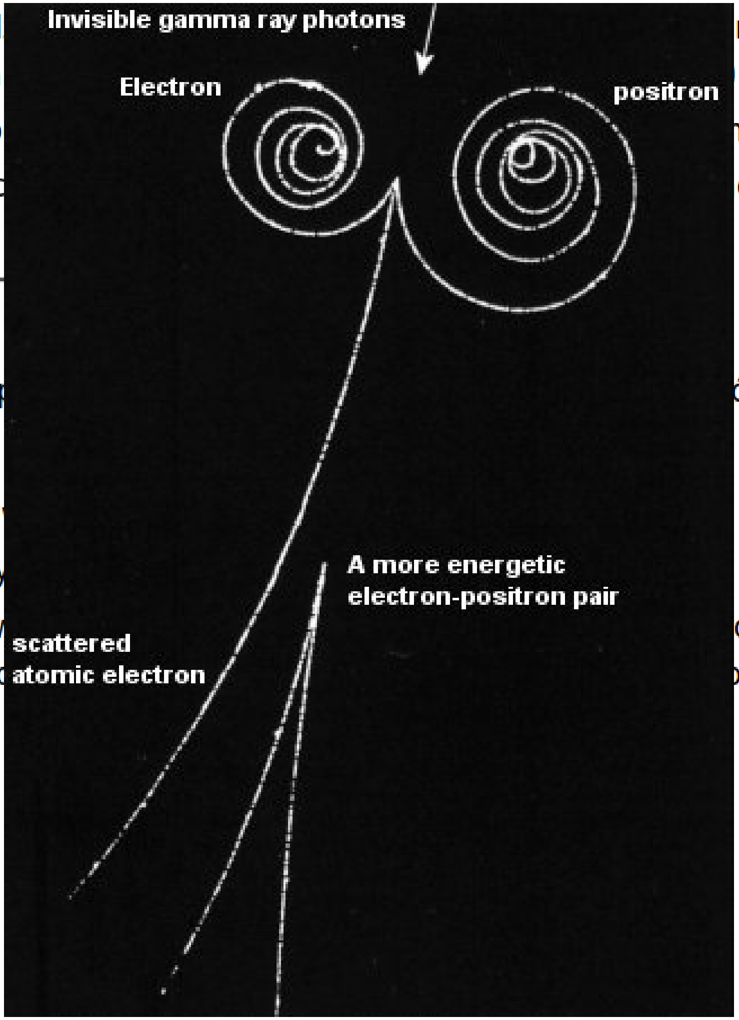
Weryfikacja hipotez statystycznych (5 pkt.)

- Przeprowadzono eksperyment naświetlania wodorowej komory pęcherzykowej wiązką fotonów w celu badania oddziaływań fotonów z protonami. Fotony powodują powstawanie par elektron-pozyton, które mogą być wykorzystane do monitorowania wiązki fotonów. Częstość występowania zdjęć z 0,1,2,... parami elektron-pozyton powinna podlegać rozkładowi Poissona. Należy wczytać dane z pliku (w pierwszej kolumnie znajduje się liczba par elektronowych na zdjęciu k, a w drugiej liczba zdjęć zawierających k par elektronowych). Wid … na - próbujemy zatem obliczyć estymator największej wiarygodności dla … slajd 13) (1 pkt.)
- Narysować na jednym wykresie punkty p … największej wiarygodności).
- Sprawdzić jakość dopasowania za pomoc … ć funkcję obliczającą statystykę testową χ2 zgodnie z wzorem $T = \sum_{k} \frac{(n_k \ldots}{\ldots}$

gdzie: nk - liczba obserwacji w k-tym binie, np … ów w k-tym binie

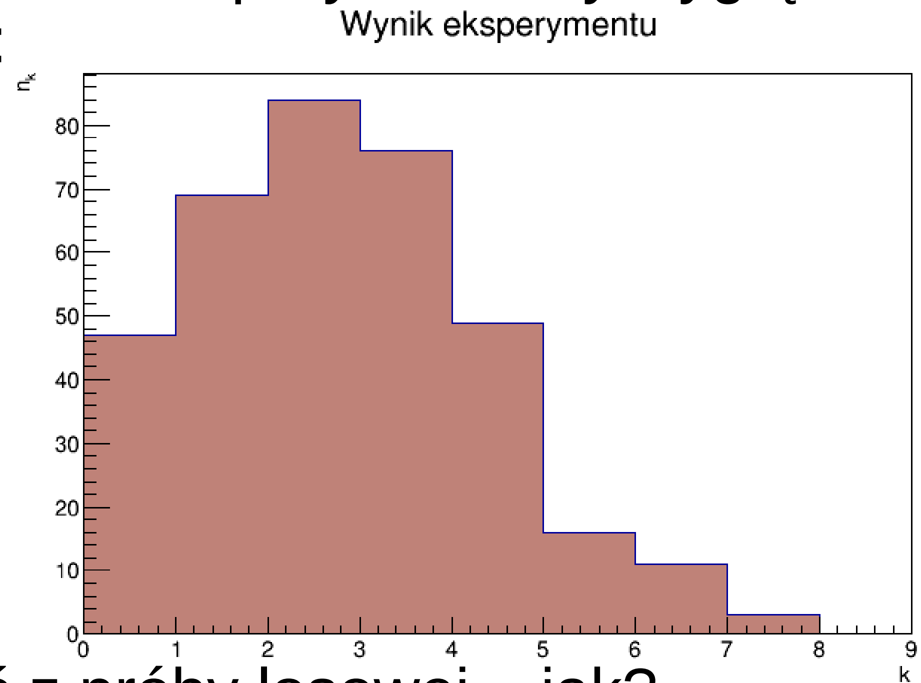
- Określić liczbę stopni swobody i obliczyć …
- Zaimplementować funkcję zwracającą wy …

Wykorzystując zaimplementowaną funkcję zw … owe podlegają rozkładowi Poissona. Dobrać odpowiednią wartość poziomu istotno … onej na ostatnich zajęciach dystrybuanty. (2 pkt.)



Test χ^2 - przykład

- Po wczytaniu danych z pliku histogram eksperymentalny wygląda następująco (nasza próba losowa):



- **Zakładamy hipotezę:** teoria mówi to jest rozkład Poissona ("na oko" zresztą tak wygląda)
- Rozkład Poissona ma tylko jeden parametr (wartość średnią):

$$f(k)=\frac{\lambda^k}{k!}e^{-\lambda}$$

- Musimy go zatem jakoś wyznaczyć z próby losowej – jak? **Na przykład metodą największej wiarygodności** – szukamy estymatora nieobciążonego największej wiarygodności o minimalnej wariancji – wyprowadziliśmy go sobie na Wykładzie 10:

$$\frac{dl}{d\lambda}=l'=\sum_{j=1}^{N}\left\{\frac{k^{(j)}}{\lambda}-1\right\}=$$

$$\frac{1}{\lambda}\sum_{j=1}^{N}\left\{k^{(j)}-\lambda\right\}=\boxed{\frac{N}{\lambda}(\bar{K}-\lambda)}$$

$$\boxed{\tilde{\lambda}=\bar{K},\ \ \sigma^2(\bar{K})=\frac{N}{\lambda}}$$

Przypomnienie – definicja estymatora o min. wariancji:

$$\boxed{\begin{gathered} l'=A(\lambda)(\tilde{\lambda}-\lambda)\\ \sigma^2(S)=\frac{1}{|A(\lambda)|}\end{gathered}}$$

Test χ^2 - przykład

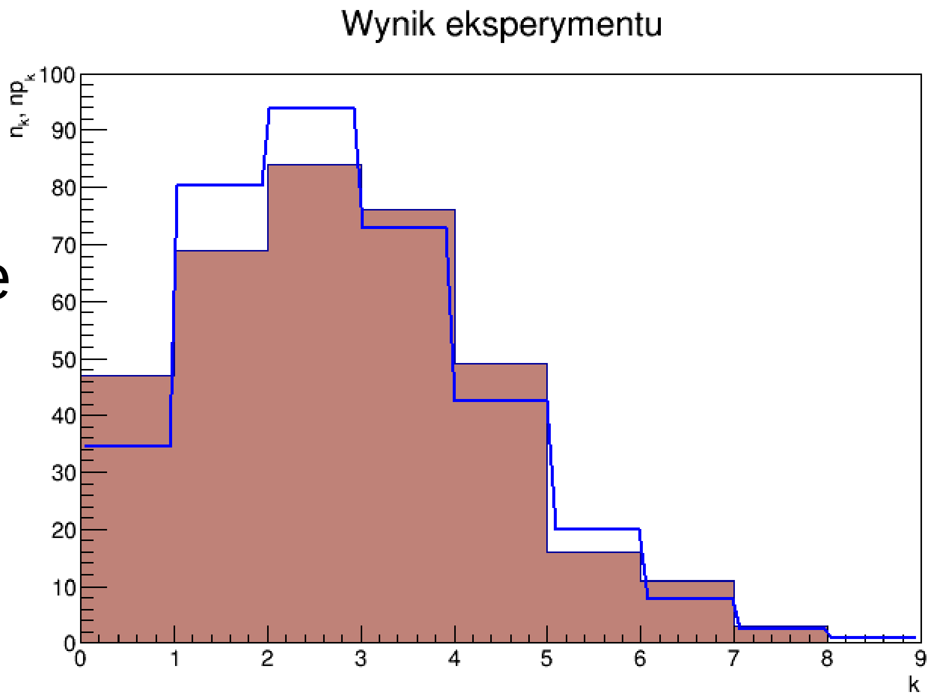
- Czyli estymatorem największej wiarygodności o minimalnej wariancji dla rozkładu Poissona jest średnia arytmetyczna z próby
- Oczywiście w naszym przypadku mamy histogram, który zawiera jakąś całkowitą liczbę wejść (całka z histogramu nie jest równa 1), wobec tego do średniej dodajemy wagi w postaci liczby wejść w danym binie i średnia staje się średnią ważoną:

$$\tilde{\lambda} = \frac{\sum_k k \cdot n_k}{\sum_k n_k}$$

- W naszym przypadku wartość ta wynosi mniej więcej: $\tilde{\lambda} \approx 2{,}33$
- Rysujemy więc funkcję:

$$n \cdot p_k = n \cdot f(k) = n \cdot \frac{\tilde{\lambda}^k}{k!} e^{-\tilde{\lambda}}, \text{gdzie } n = \sum_k^7 n_k$$

- Jak teraz sprawdzić, czy faktycznie nasza hipoteza jest słuszna?
- **Testujemy dobroć dopasowania**



Test χ^2 - przykład

- Musimy zatem wyznaczyć wartość statystyki testowej T:

$$T=\sum_{k=0}^{7} u_i^2=\sum_{k=0}^{7} \frac{(n_k-n\,p_k)^2}{np_k}\approx 10{,}53$$

- Co dalej? Zakładamy poziom istotności, na przykład: $\alpha=0{,}01$
- Musimy jeszcze określić liczbę stopni swobody – ile ich jest?
  - liczba binów (8) minus 1 minus liczba parametrów (1)

$$r-1-p=8-1-1=6$$

- Teraz szukamy odpowiedniego kwantyla rozkładu χ^2 o 6 stopniach swobody: $\chi^2_{1-\alpha}=\chi^2_{0,99}\approx 16{,}81$
- Porównujemy statystykę z kwantylem: $T=10{,}51<\chi^2_{0,99}=16{,}81$
- **Warunek** $T>\chi^2_{1-\alpha}$ **nie jest spełniony, zatem na poziomie istotności** $\alpha=0{,}01$ **nie ma podstaw do odrzucenia hipotezy**

Test równości wariancji (test F-Fischera)

Test równości wariancji (F-Fischera)

- **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

$$F=s_1^2/s_2^2 \qquad s^2(X)=\frac{1}{N-1}\sum_{i=1}^{N}\left(X_i-\bar{X}\right)^2 \qquad s^2(\bar{X})=\frac{1}{N\cdot(N-1)}\sum_{i=1}^{N}\left(X_i-\bar{X}\right)^2=s^2(X)/N$$

- Jeśli hipoteza o równości wariancji jest **prawdziwa**, to iloraz F powinien być bliski jedności $F\sim 1$

Test równości wariancji (F-Fishera)

- Można udowodnić, że tego typu wielkość ma rozkład F-Fischera

$$f(F)=\left(\frac{f_1}{f_2}\right)^{\frac{1}{2}f_1}\frac{\Gamma\left(\frac{1}{2}(f_1+f_2)\right)}{\Gamma\left(\frac{1}{2}f_1\right)\Gamma\left(\frac{1}{2}f_2\right)}F^{\frac{1}{2}f_1-1}\left(1+\frac{f_1}{f_2}F\right)^{-\frac{1}{2}(f_1+f_2)} \qquad f_1=N_1-1;\ \ f_2=N_2-1$$

- Szukamy zatem analogicznie wartości granicznej określającej obszar krytyczny, która jest odpowiednim kwantylem rozkładu F-Fischera:

$$P\left(\frac{s_1^2}{s_2^2}>F_{1-\alpha}\right)=\alpha$$

- Ostatecznie sprawdzamy zatem warunek: $\boxed{\frac{s_1^2}{s_2^2}>F_{1-\alpha}}$
- **To jest test jednostronny**, na ogół posługujemy jednak **testem dwustronnym**:

$$\boxed{\frac{s_g^2}{s_k^2}>F_{1-\alpha/2}(f_g,f_k),\ \text{gdzie f to liczby stopni swobody}}$$

Test równości wariancji (F-Fischera)

- Czyli w praktyce musimy zweryfikować hipotezę:

$$\frac{s_g^2}{s_k^2} > F_{1-\alpha/2}(f_g, f_k) = F''_{\alpha}(f_g, f_k)$$

- 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 **odrzucić**

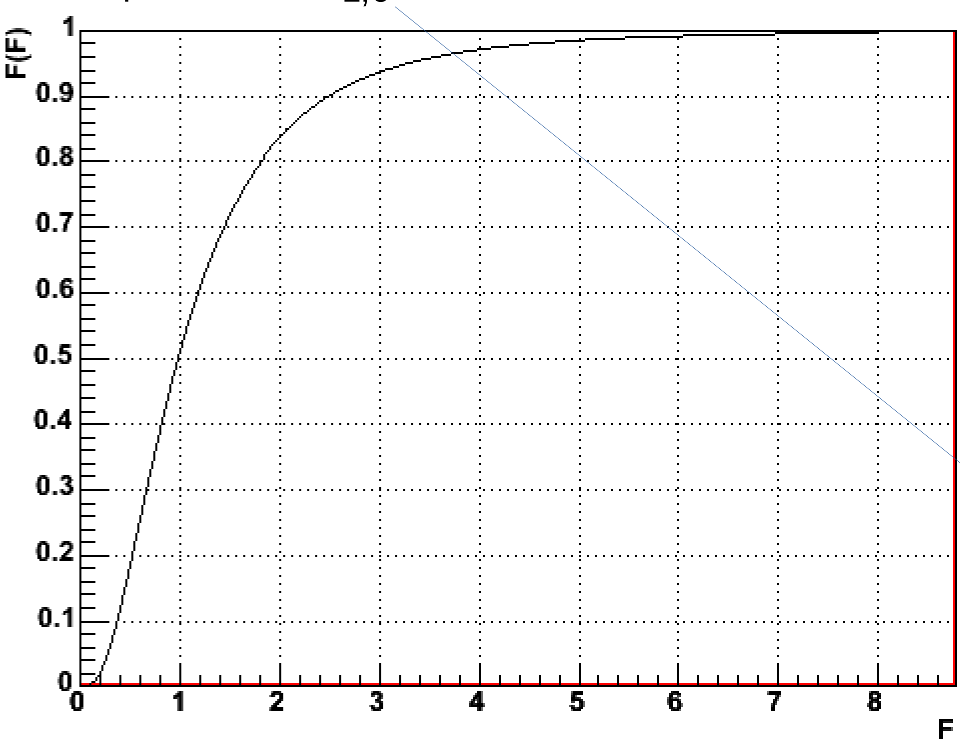
Test równości wariancji – przykład

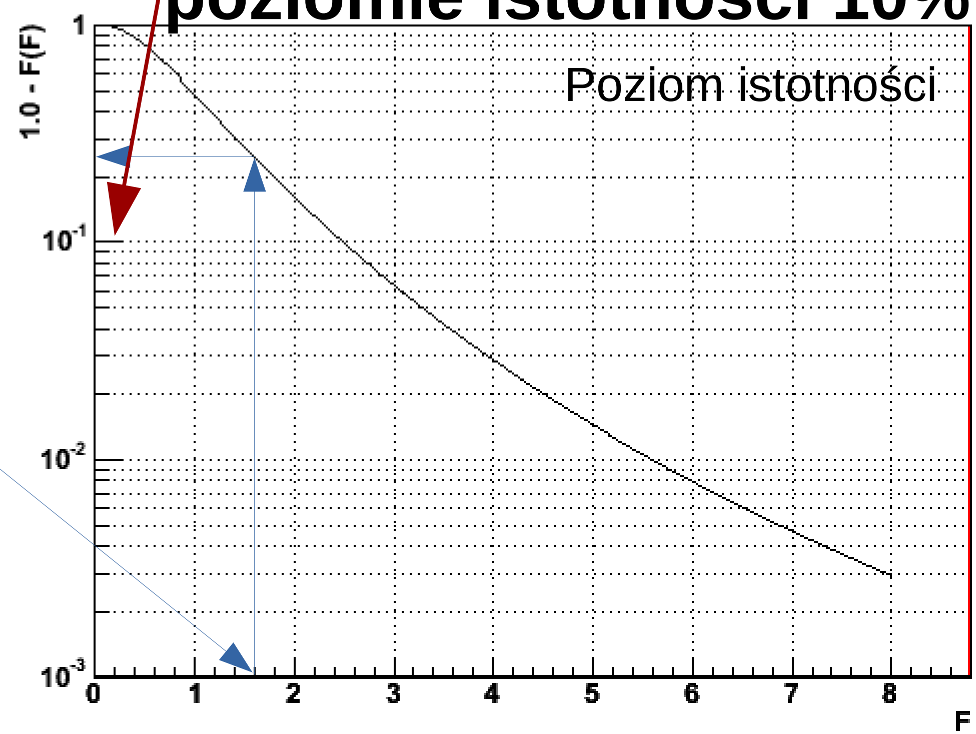
Numer pomiaru	Przyrząd 1	Przyrząd 2				
1	100	97	0,2	0,04	-2,86	8,16
2	101	101	1,2	1,44	1,14	1,31
3	102	102	2,2	4,84	2,14	4,59
4	100	99	0,2	0,04	-0,86	0,73
5	98	101	-1,8	3,24	1,14	1,31
6	97	98	-2,8	7,84	-1,86	3,45
7	100	101	0,2	0,04	1,14	1,31
8	101		1,2	1,44		
9	99		-0,8	0,64		
10	100		0,2	0,04		
Średnia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

- Korzystamy z kwantyli funkcji F

$$F''_{0,2}(6,9)=F_{0,9}(6,9)=2.51$$
$$\boxed{F''_{0,1}(6,9)=F_{0,95}(6,9)=3.29}$$
$$F''_{0,02}(6,9)=F_{0,99}(6,9)=5.61$$
$$F''_{0,01}(6,9)=F_{0,995}(6,9)=6.89$$

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





Porównanie wartości średnich
(test t-Studenta)

Test równości średnich (t-Studenta)

- **Problem:** porównywanie wartości średnich dwóch prób losowych
- **Przykład:** badamy średni wzrost studentek 1 roku w Warszawie (populacja X) oraz w Nowym Jorku (populacja Y)
- **Pytanie (hipoteza):** czy wartości średnie obu populacji, na podstawie pobranych prób losowych, są jednakowe?
- Tak postawiona hipoteza cicho zakłada, że X i Y to te same populacje
- Powyższe rozważania możemy **uogólnić** na porównanie wartości średnich dwóch prób losowych z populacji X oraz Y o liczebnościach N_1 i N_2

Zastosowanie testu t-Studenta

- **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

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:

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

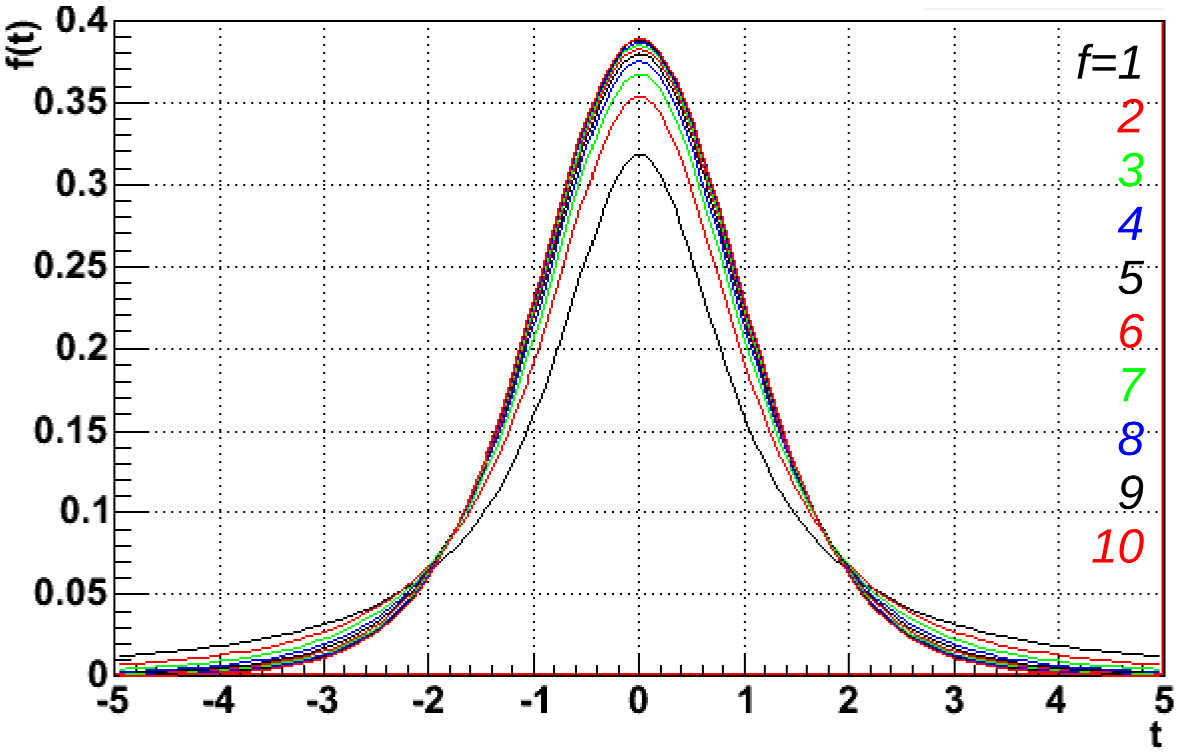
- Można udowodnić, że zmienna $\Delta/s(\Delta)$ podlega rozkładowi t-Studenta z liczbą stopni swobody $f=N_1+N_2-2$

- 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



- 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 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)

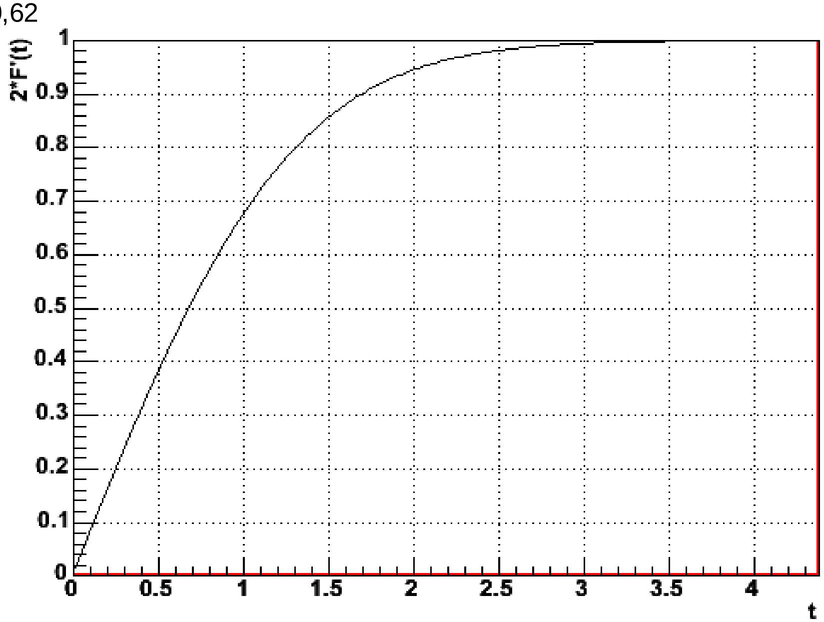
Test różnic t-Studenta - przykład

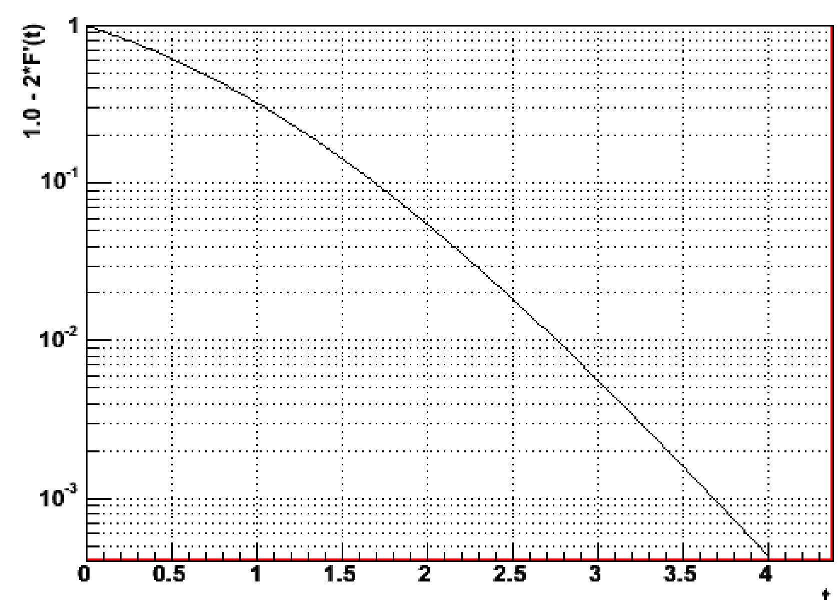
Numer pomiaru	Przyrząd 1	Przyrząd 2				
1	**100**	**97**	-0,21	0,04	-3,8	14,44
2	**101**	**101**	0,79	0,62	0,2	0,04
3	**102**	**102**	1,79	3,2	1,2	1,44
4	**100**	**99**	-0,21	0,04	-1,8	3,24
5	**98**	**101**	-2,21	4,89	0,2	0,04
6	**97**	**108**	-3,21	10,31	7,2	51,84
7	**100**	**101**	-0,21	0,04	0,2	0,04
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**






WYDZIAŁ FIZYKI
POLITECHNIKA WARSZAWSKA
POLITECHNIKA WARSZAWSKA
WARSAW UNIVERSITY OF TECHNOLOGY

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 $(N-1)s_X^2 = fs_X^2$ 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 $f \rightarrow \infty$
- 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 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**