Komputerowa analiza danych doświadczalnych

Wykład podsumowujący
21.05.2021

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

Semestr letni 2020/2021

Parametry rozkładów prawdopodobieństwa

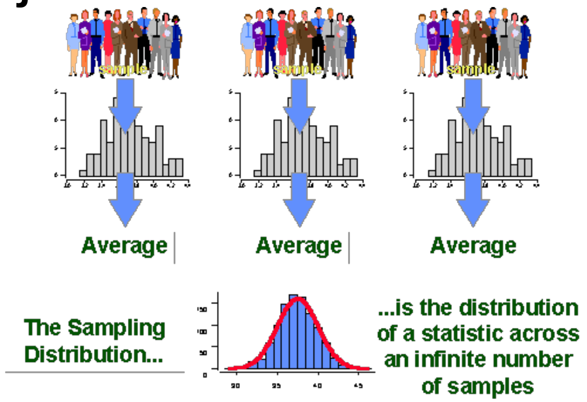
Generowanie liczb i metody Monte Carlo

Dopasowanie modelu do danych

Testy statystyczne

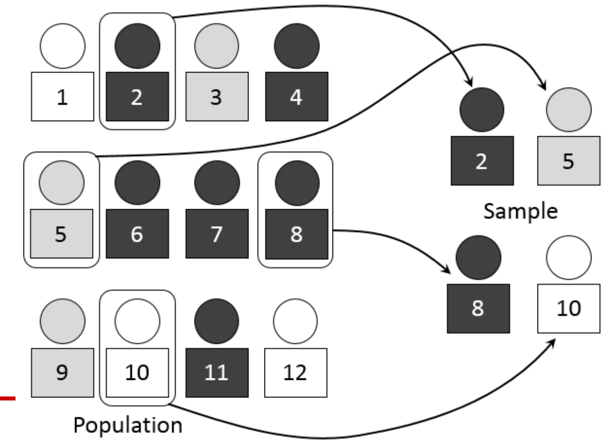
Statystyczna analiza danych

- Statystyczna analiza danych:
  - traktujemy pomiar jako pewien element zbioru wszystkich możliwych pomiarów (pewnej cechy **populacji** o danym rozkładzie prawdopodobieństwa – najczęściej nieznanym)
  - na podstawie skończonej liczby pomiarów, obserwacji (**próby losowej**, podzbioru populacji), która ma swój rozkład prawdopodobieństwa (znany z pomiarów czy obserwacji), próbujemy dowiedzieć się czegoś (czyli **estymować**) na temat parametrów rozkładu całej populacji
  - innymi słowy, na podstawie próby losowej (pomiarów, obserwacji) stawiamy hipotezy i wyciągamy wnioski dotyczące interesującej nas cechy całej populacji



https://www.proprofs.com/quiz-school/story.php?title=3dq-sampling-distributions

https://en.wikipedia.org/wiki/Sample_(statistics)



Rozkłady 1D - momenty

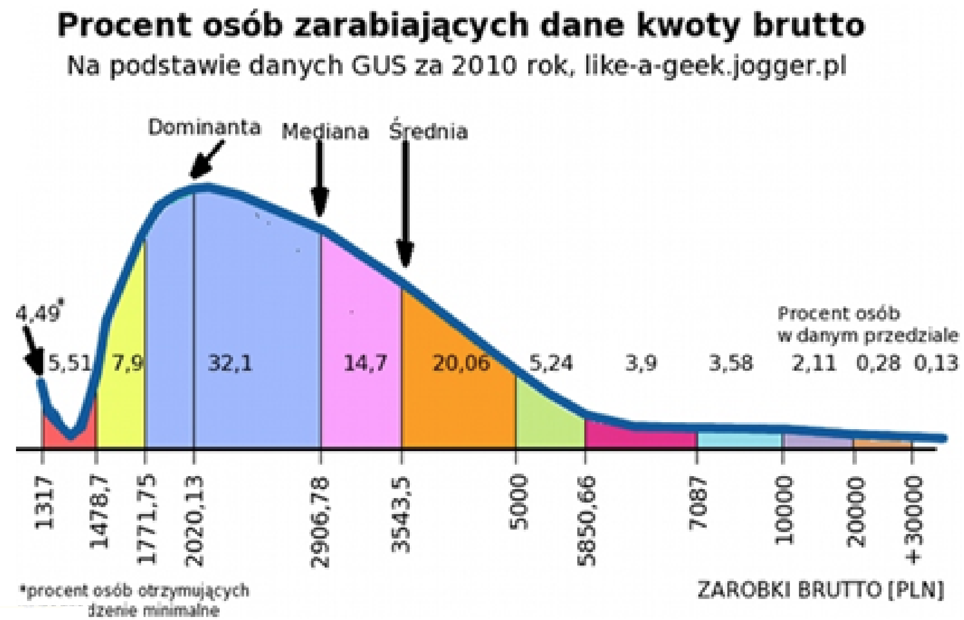
$$m_l = E\left(X^l\right) = \int_{-\infty}^{\infty} x^l f(x)\,dx$$ momenty zwykłe

$$m_1 = \mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

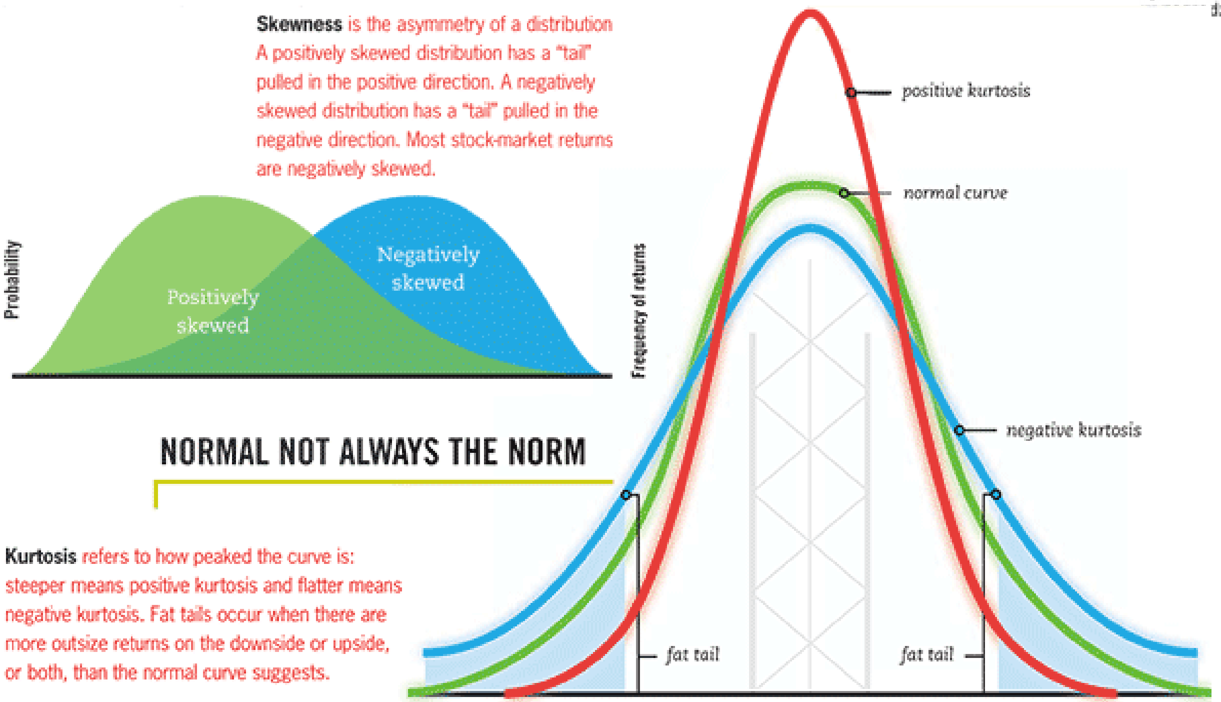
$$\mu_l = E\left(\left(X - \hat{x}\right)^l\right)$$ momenty centralne

$$\mu_2 \equiv \sigma^2(X) \equiv E\left(\left(X-\hat{x}\right)^2\right) = \int_{-\infty}^{\infty} \left(x-\hat{x}\right)^2 f(x)\,dx$$

$$\gamma = \frac{\mu_3}{\sigma^3} \qquad K = \frac{\mu_4}{\sigma^4} - 3$$ wsp. asymetrii, kurtoza



http://lh3.ggpht.com/-UhjcSGuME9Q/UgCqCj00_nI/AAAAAAAAWXU/-0ZIMA9pPnU/image_thumb%25255B2%25255D.png?imgmax=800



http://www.advisor.ca/wp-content/uploads/2012/07/normal-not-always-the-norm.gif

dominanta

$$P\left(X = x_{max}\right) = max$$

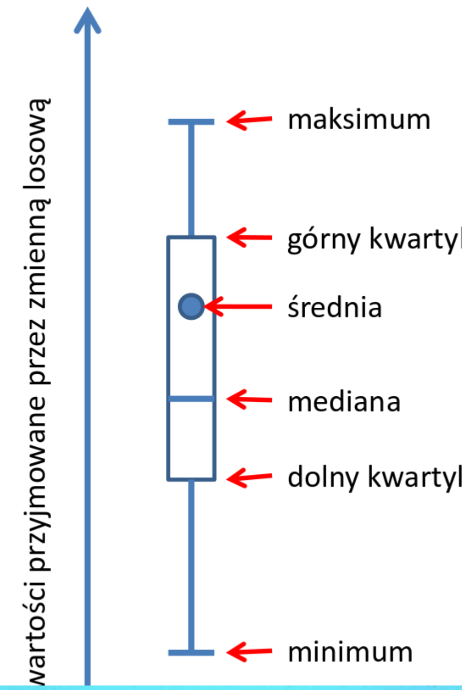
$$\frac{df(x)}{dx} = 0 \qquad \frac{d^2 f(x)}{dx^2} < 0$$

Momenty to **uśrednienia** danych podniesione do kolejnych **potęg**

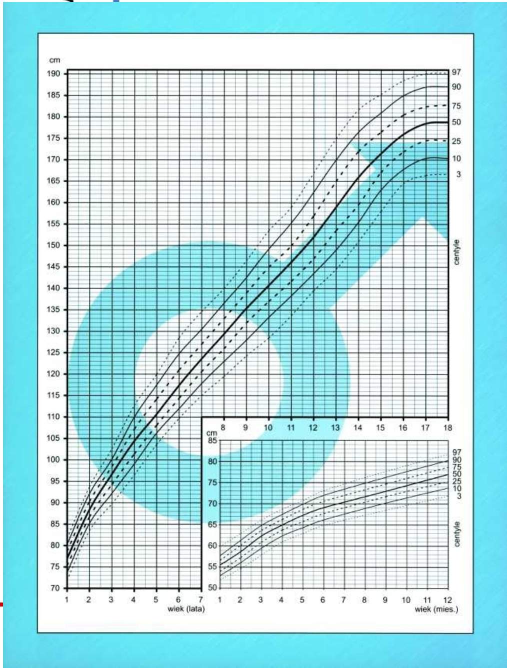
Rozkłady 1D – kwantyle

- **Mediana** dzieli rozkład prawdopodobieństwa na dwa obszary o równym prawdopodobieństwie $F(x_{0,5}) = P(X < x_{0,5}) = 0,5$
- Mediana $x_{0,5}$ jest **kwantylem** (*ang. quantile*) rzędu 0,5
- Ogólna definicja **kwantylu rzędu q**, x_q: $F(x_q) = P(X < x_q) = q$
  - **kwartyl** dolny $x_{0,25}$
  - kwartyl górny $x_{0,75}$
  - **decyle** $x_{0,1}, x_{0,2}, \ldots, x_{0,9}$
  - funkcja $x_q(q)$ jest funkcją <u>odwrotną</u> do dystrybuanty

$$F(x_q) = \int_{-\infty}^{x_q} f(x)\,dx = q, q \in \langle -1;1 \rangle$$



- Kwantyl rzędu q jest taką liczbą x_q, że $q \cdot 100\%$ elementów w danej próbce (populacji) ma wartość pomiaru (badanej cechy) nie większą niż x_q
- W przypadku ogonów rozkładu kwantyle mogą być lepszą wielkością niż momenty
- **Momenty i kwantyle to dwa najczęstsze opisy numeryczne danych liczbowych**



Rozkłady 2D

momenty

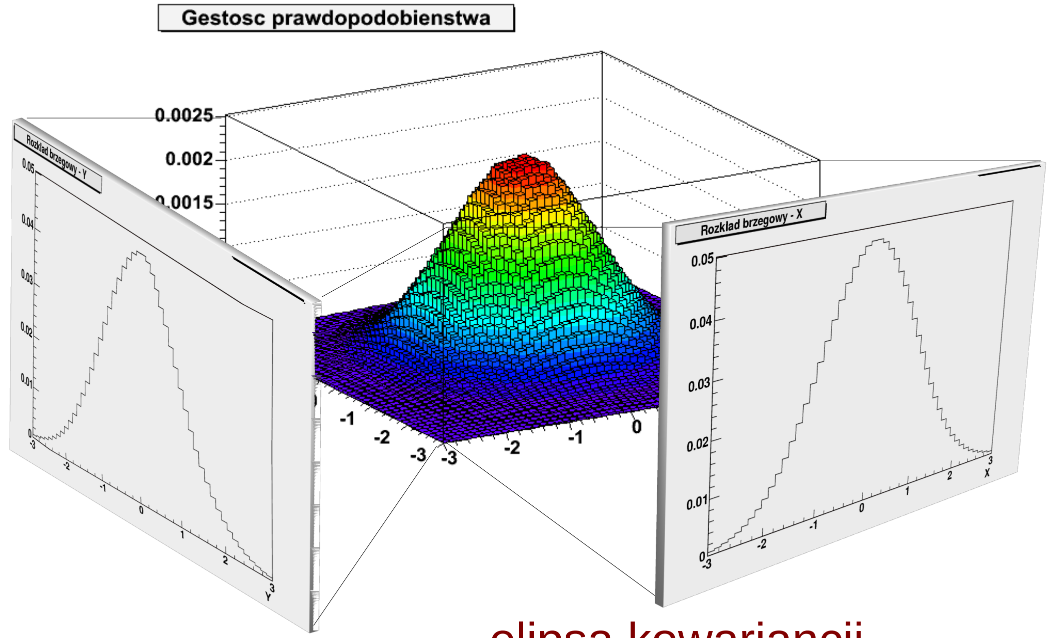
$$\lambda_{10}=E(x)=\hat{x}$$

$$\lambda_{01}=E(y)=\hat{y}$$

$$\mu_{11}=E((X-\hat{x})(Y-\hat{y}))=cov(X,Y)$$

$$\mu_{20}=E\left((X-\hat{x})^2\right)=\sigma^2(X)$$

$$\mu_{02}=E\left((Y-\hat{y})^2\right)=\sigma^2(Y)$$



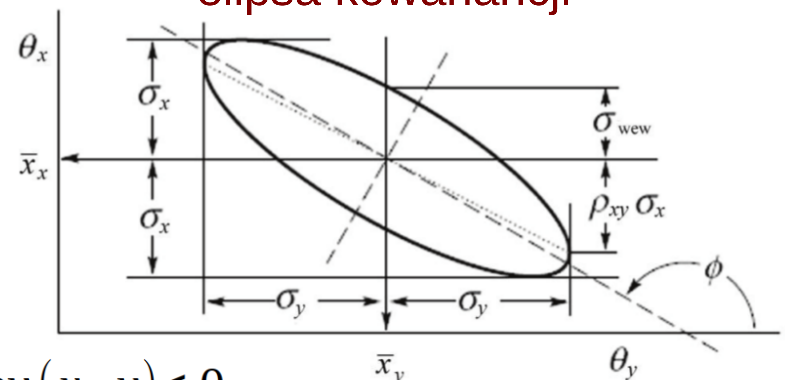
kowariancja

$$cov(X,Y)=\mu_{11}=E(X\cdot Y)-E(X)\cdot E(Y)$$

wsp. korelacji (Pearsona)

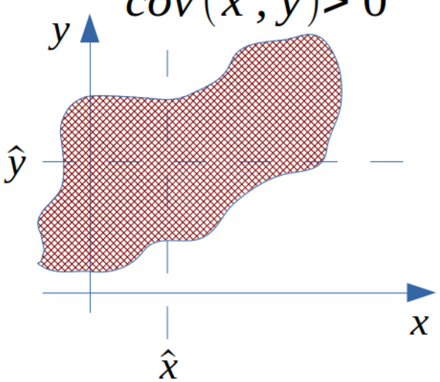
$$\rho(X,Y)=\frac{cov(X,Y)}{\sigma(X)\sigma(Y)} \qquad -1\leqslant\rho(X,Y)\leqslant 1$$

elipsa kowariancji

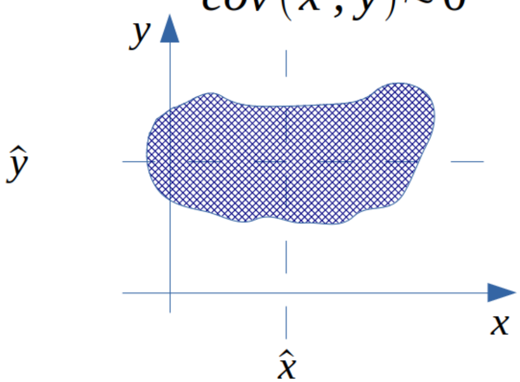


Współczynnik korelacji a liniowa zależność zmiennych
(niekoniecznie **przyczyna-skutek**)

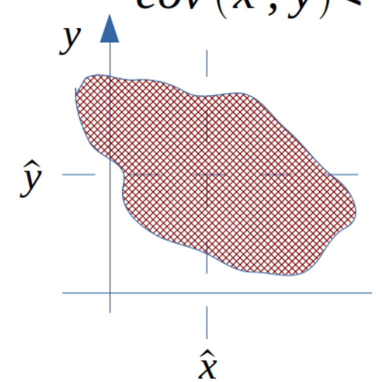
$$cov(x,y)>0$$

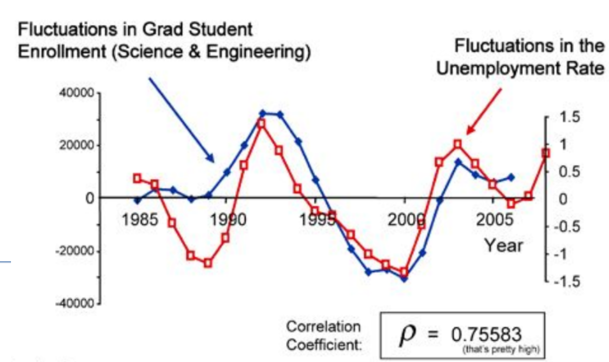


$$cov(x,y)\approx 0$$



$$cov(x,y)<0$$

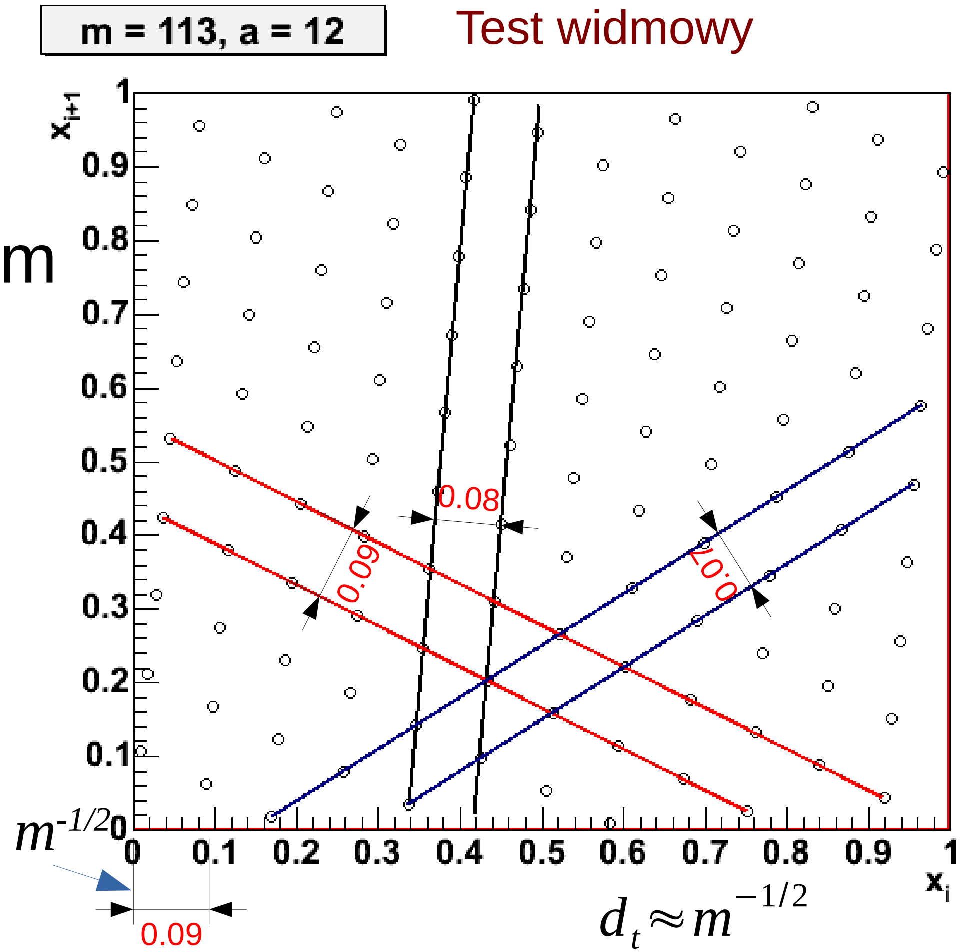




Generatory liczb pseudolosowych

- Komputer, urządzenie determinisityczne, może generować tylko liczby pseudolosowe
  - kolejna generowana liczba jest funkcją liczb wcześniej wygenerowanych
  - *ziarno* (*seed*) – wartość początkowa (można ją ustalić np. z zegara systemowego)
- Generujemy liczby z jednakowym prawdopodbieństwem (rozkład jednorodny)

Test widmowy



Liniowy kongruentny generator liniowy (LCG)

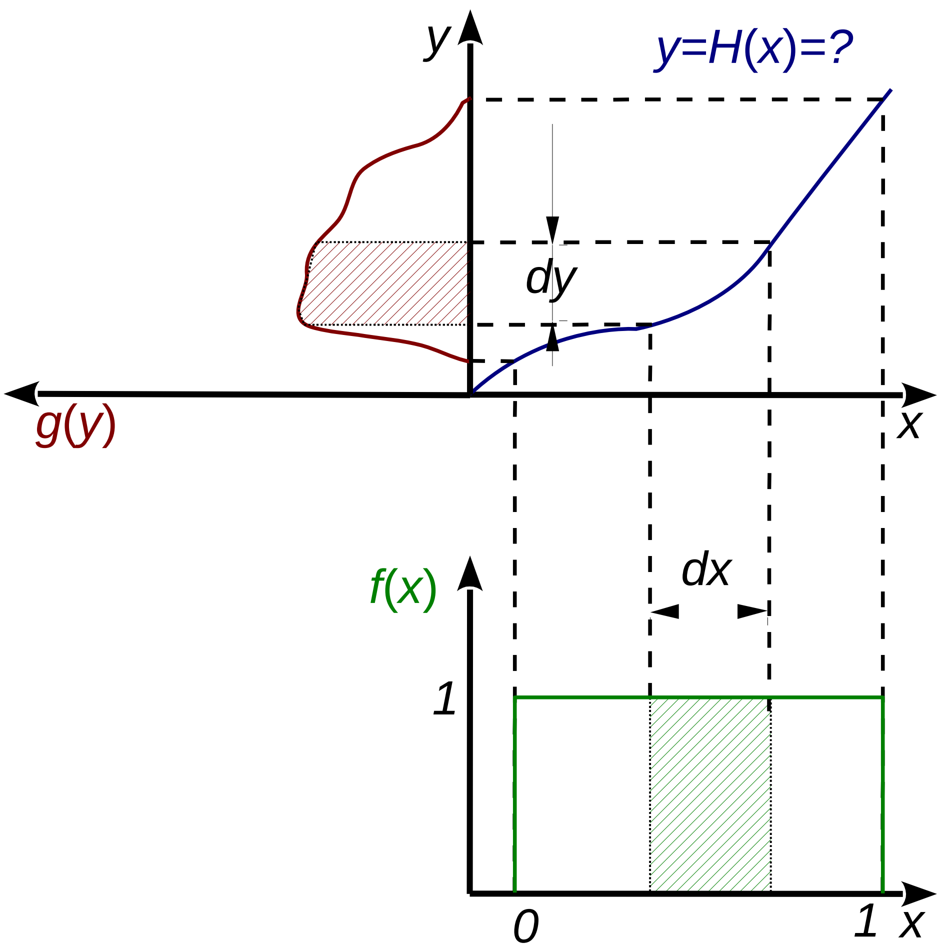
$$x_{j+1} = (a \cdot x_j + c) \, mod \, m$$

Multiplikaywny liniowy kongruentny generator liniowy (MLCG)

$$x_{j+1} = (a \cdot x_j) \, mod \, m$$

Transformacja rozkładu jednorodnego

- **Metoda odwrotnej dystrybuanty**
- Transformację rozkładu jednostajnego możemy wykorzystać do generowania liczb losowych o skomplikowanych gęstościach prawdopodobieństwa



$$f(x)dx = g(y)dy$$

$$\text{gdy} \quad f(x) \equiv 1 \Rightarrow dx = g(y)dy = dG(y)$$

dystrybuanta

$$g(y) = G'(y)$$

$$\int dx = \int dG(y)$$

$$x = G(y)$$

musi istnieć

$$\boxed{y = G^{-1}(x) \equiv H(x)}$$

$$x_{min} = G(y_{min}),\ x_{max} = G(y_{max})$$

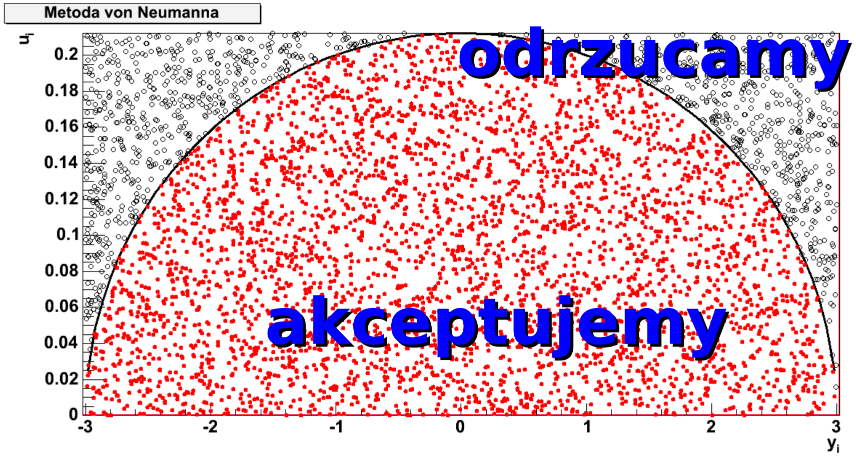
$$y_{min} = G^{-1}(0),\ y_{max} = G^{-1}(1)$$

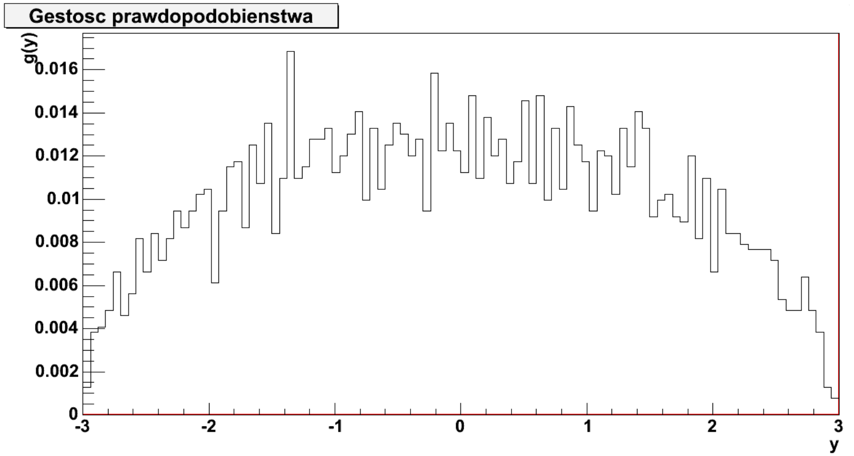
Czyli: liczymy dystrybuantę $G(y)$
a następnie funkcję odwrotną $G^{-1}(y)$

Zmienna losowa X po transformacji $y=G^{-1}(x)$ ma rozkład $g(y)$

Metoda (akceptacji) von Neumanna

- Metoda generacji liczb metodą odwrotnej dystrybuanty ma swoje ograniczenia – dystrybuanta musi być znana (I być funkcją wzajemnie jednoznaczną), czyli musi istnieć funkcja odwrotna
- Metoda (akceptcji-odrzuceń) von Neumanna wymaga znajomości jedynie gęstości prawdopodobieństwa i pozwala na otrzymanie liczb z praktycznie dowolnego rozkładu
- Jak to działa?
  – generujemy parę liczb z rozkładu jednorodnego: (y_i, u_i)
  – sprawdzamy, czy $u_i < g(y_i)$
  – jeśli warunek jest spełniony, <u>akceptujemy</u> liczbę y_i,
  jeśli nie - odrzucamy

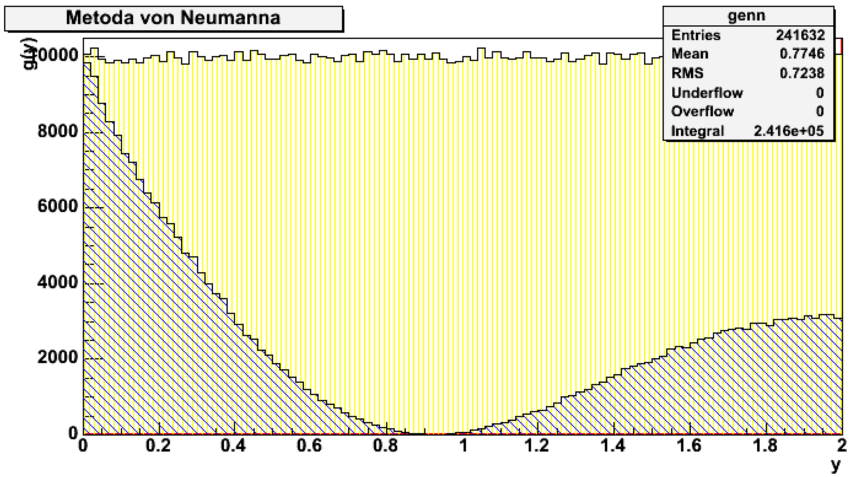


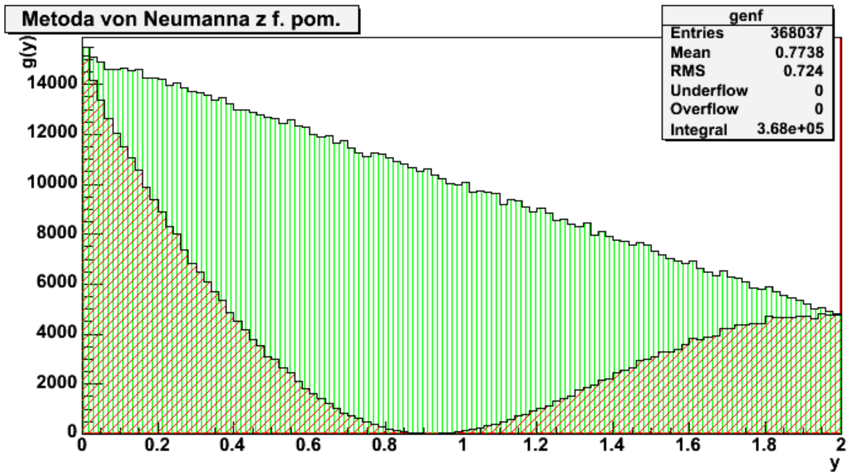


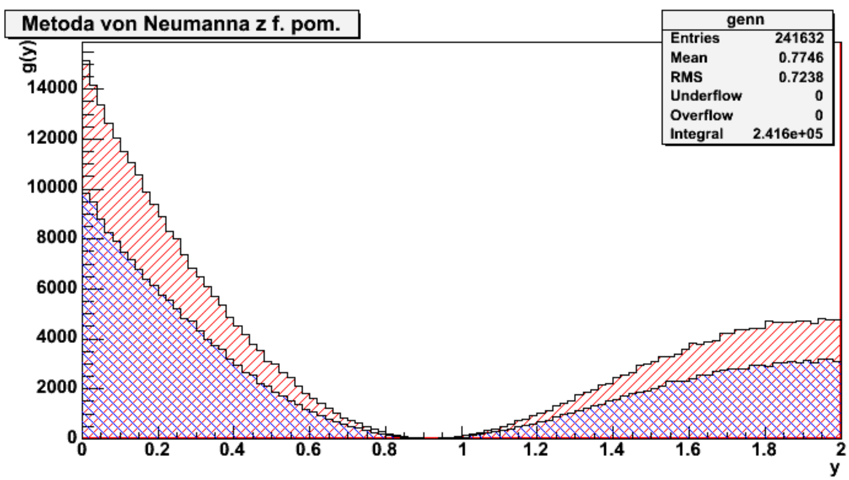
Metoda von Neumanna z funkcją pom.

- Wydajność metody von Neumanna można poprawić, jeśli odpowiednio zawężymy obszar losowania:
  - wprowadzamy funkcję pomocniczą $s(y)$, z ktorej “łatwo” wygenerować zmienne losowe (np. metodą odwrotnej dystrybuanty), i która spełnia warunek: $g(y) \leqslant c \cdot s(y), \quad a < y < b$
  - generujemy liczbę losową y_i z rozkładu $s(y)$ na przedziale $a < y_i < b$ oraz liczbę u_i z rozkładu jednorodnego na przedziale $0 < u_i < 1$
  - <u>odrzucamy</u> liczbę y_i, jeżeli: $u_i \geqslant \dfrac{g(y_i)}{c \cdot s(y_i)}$
  - wydajność metody:

$$E = \frac{\int_a^b g(y)\,dy}{c\int_a^b s(y)\,dy}$$

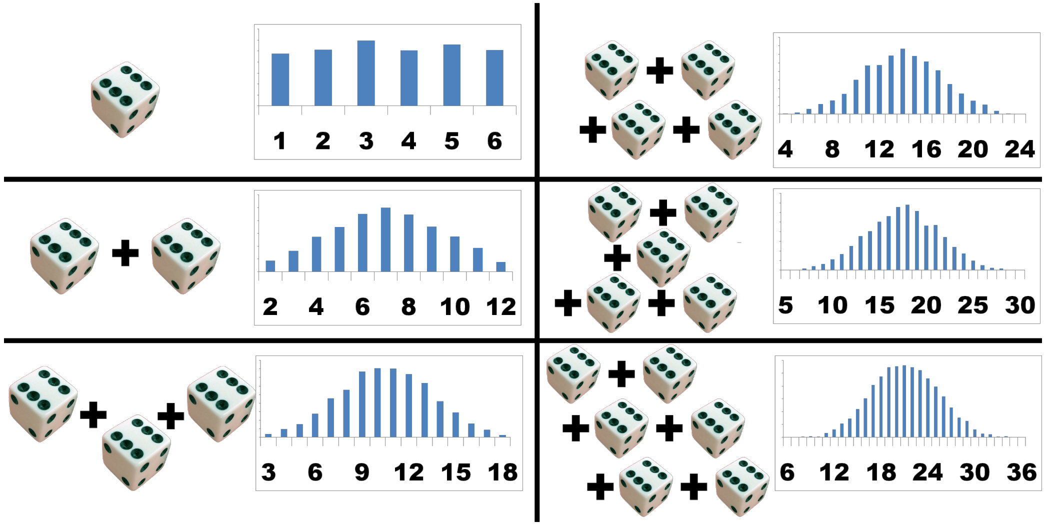


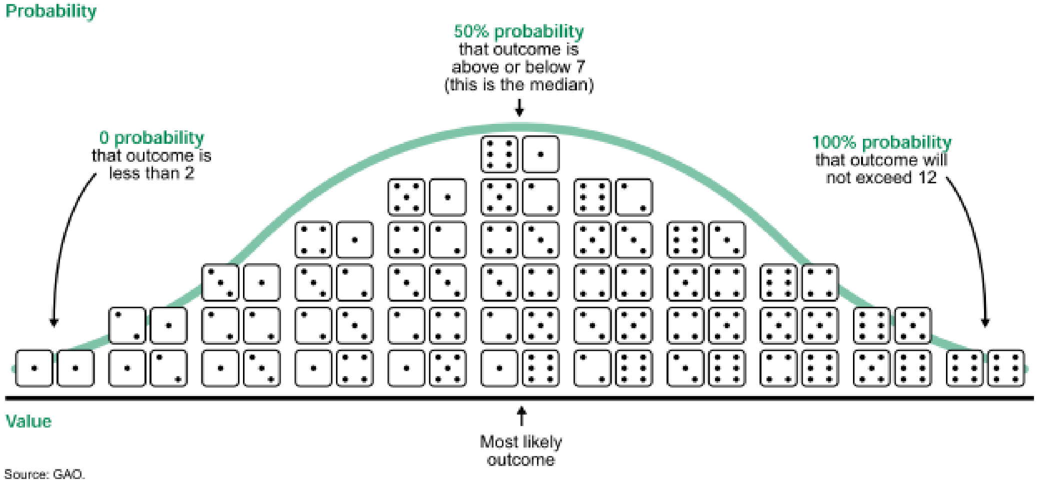




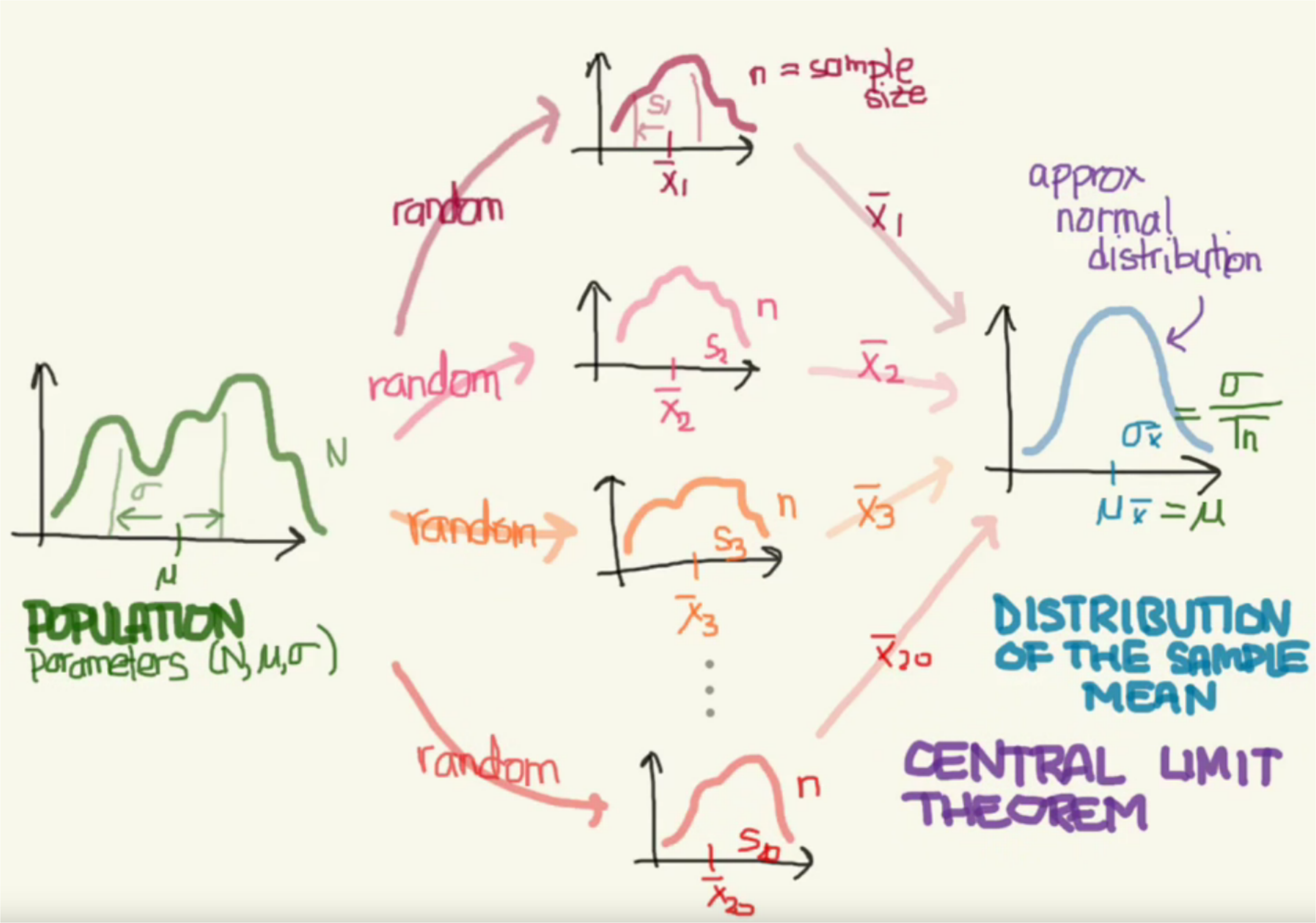
Centralne twierdzenie graniczne

- Wyobraźmy sobie eksperyment polegający na rzucie kostką (kostkami) i obserwowaniu całkowitej liczby oczek:
  - kolejne rzuty kostką (kostkami) są niezależne
  - jeśli rzucamy kostką jednokrotnie (albo 1 kostką), to prawdopodobieństwo uzyskania danej wartości jest jednakowe
  - jeśli rzucamy kostką dwukrotnie (albo 2 kostkami), to prawdopodobieństwo uzyskania sumy oczek nie jest już jednakowe
  - jeśli rzucimy kostką n-krotnie (n-kostkami) $\rightarrow$ rozkład normalny



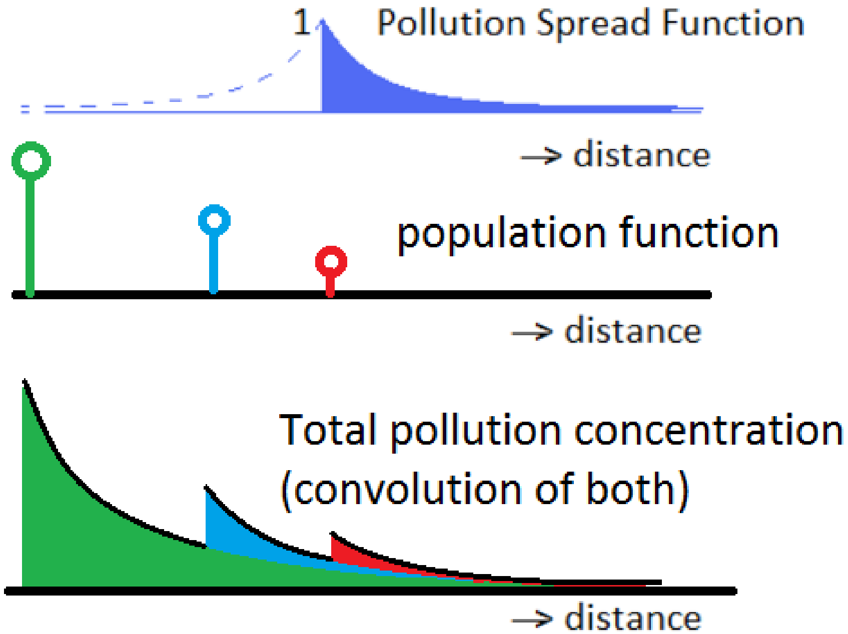


Centralne twierdzenie graniczne



Suma zmiennych losowych jako splot

https://www.quora.com/The-density-function-of-the-sum-of-two-random-variables-is-the-convolution-of-their-respective-densities-What-is-the-intuition-behind-this



- Wyobraźmy sobie taką sytuację:
  - Mieszkasz w wiosce obok rzeki
  - Mieszkańcy wioski wrzucają do rzeki odpady biologiczne
  - Koncentracja odpadów w funkcji odległości od miejsca zrzutu (*Pollution Spread Function, PSF*) jest zależna od ich rozkładu przez mikroorganizmy w rzece
  - Ilość wrzucanych odpadów zależy od populacji miejscowości na rzece
- Jaka jest pełna funkcja opisująca poziom zanieczyszczeń w rzece?
- Jest to **splot** dwóch rozkładów – funkcji populacji oraz funkcji koncentracji odpadów
- Innymi słowy, zastępujemy każdy punkt w funkcji populacji przez funkcję koncentracji przeskalowaną przez wagę funkcji populacji

Suma zmiennych losowych jako splot

- Przypadek splotu dwóch rozkładów jednorodnych:

$$f_x(x) = \begin{cases} 1, & 0 \leqslant x < 1 \\ 0, & \text{w przeciwnym razie} \end{cases} \qquad f_y(y) = \begin{cases} 1, & 0 \leqslant y < 1 \\ 0, & \text{w przeciwnym razie} \end{cases}$$

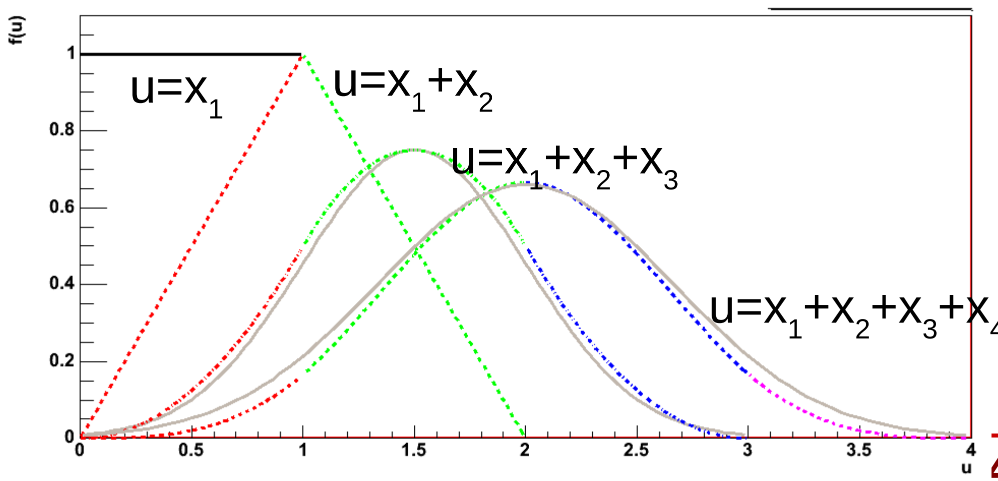
$$f(u) = \int_0^1 f_x(x) f_y(u-x)\,dx = \int_0^1 f_y(u-x) \quad \begin{matrix} v = u - x \\ dv = -dx \end{matrix} \Rightarrow \quad f(u) = -\int_u^{u-1} f_y(v)\,dv = \int_{u-1}^{u} f_y(v)\,dv$$

- Zmienna u zmienia się od 0 do 2, zatem rozważmy 2 przypadki:

$$(a)\quad 0 \leqslant u < 1: \quad f_1(u) = \int_0^u f_y(v)\,dv = \int_0^u dv = u$$

$$(b)\quad 1 \leqslant u < 2: \quad f_2(u) = \int_{u-1}^{1} f_y(v)\,dv = \int_{u-1}^{1} dv = 2 - u$$

- Zgodnie z CTG – im więcej rozkładów w splocie, tym bardziej rozkład sumy przypomina rozkład Gaussa:



Estymatory

- Typowy problem analizy danych: znamy (np. z prawa fizycznego) ogólną postać gęstości prawdopodobieństwa w danej populacji, należy “jedynie” wyznaczyć parametry tego rozkładu. Przykład:
  - mierzymy rozpad radioaktywny w czasie: $N(t)=N_0(1-\exp(-\lambda t))$
  - parametr λ wyznaczamy na podstawie próby – mierząc skończoną ilość razy ilość rozpadów w czasie → wynik nigdy nie będzie dokłady, bo próba jest skończona, mamy problem **estymacji parametrów**
  - poszukiwana wielkość uzyskiwana jest funkcją elementów próby (**statystyką**) i jest nazywana **estymatorem**: $S=S(X_1,X_2,\ldots,X_n)$
  - estymator jest **nieobciążony**, jeżeli niezależnie od liczebności próby jego wartość oczekiwana jest równa wartości estymowanego parametru:

$$E\left(S(X_1,X_2,\ldots,X_n)\right)=\lambda,\ \text{dla każdego}\ n$$

  - estymator jest **zgodny**, jeżeli jego wariancja znika:

$$\lim_{n\to\infty}\sigma\left(S(X_1,X_2,\ldots,X_n)\right)=0$$

Estymator wartości oczekiwanej

Populacja

- opisana funkcją gęstości:

$$f(x) = P(X = x)$$

- posiada **wartość oczekiwaną**:

$$E(X) = \hat{x} = \int_{-\infty}^{\infty} x f(x)\, dx$$

- wartość oczekiwana rozkładu to **jedna liczba**
  - nie jest zmienną losową
  - chcemy ją zmierzyć doświadczalnie
- np. dla rozkł. Gaussa:

$$f(x) \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \mu$$

Próba losowa

- zakładmy, że **średnia arytmetyczna** z elementów próby jest estymatorem wartości oczekiwanej
- **średnia arytmetyczna** jest statystyką:

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \ldots X_n)$$

- **jest zmienną losową** (zależy od elementów próby)
- posiada swoją **wartość oczekiwaną** oraz **wariancję**
- oczekujemy, że będzie ona *estymatorem nieobciążonym* i *zgodnym* wartości oczekiwanej populacji:

$$E(\bar{X}) = E(X) = \hat{x}, \text{ dla każdego } n$$

$$\lim_{n \to \infty} \sigma(\bar{X}) = 0$$

- **Na wykładzie pokazane jest jak to sprawdzić**

Estymatory - podstawowe

- Przykładowe **estymatory nieobciążone**:
  - **wartości oczekiwanej populacji → *średnia arytmetyczna z próby:***

$$\bar{X}=\frac{1}{n}(X_1+X_2+\ldots+X_n)$$

  - **wariancji populacji → *średnia odchyleń kwadratowych:***

$$S^2(X)=\frac{1}{n-1}\left((X_1-\bar{X})^2+(X_2-\bar{X})^2+\ldots+(X_n-\bar{X})^2\right)$$

-1 wynika z jednego równania więzów (istnienie średniej)

- Wariancje (niepewności) estymatorów:
  - **wariancja** *średniej arytmetycznej*:

$$\sigma^2(\bar{X})=\frac{1}{n}S^2(X)=\frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

  - **wariancja** *średniej odchyleń kwadratowych*:

$$\sigma^2\left(S^2(X)\right)=S^4(X)\left(\frac{2}{n-1}\right)$$

- **Uwaga!** Wariancje estymatorów są również estymatorami – możemy więc liczyć np. wariancję wariancji średniej arytmetycznej, itd.

Metoda największej wiarygodności

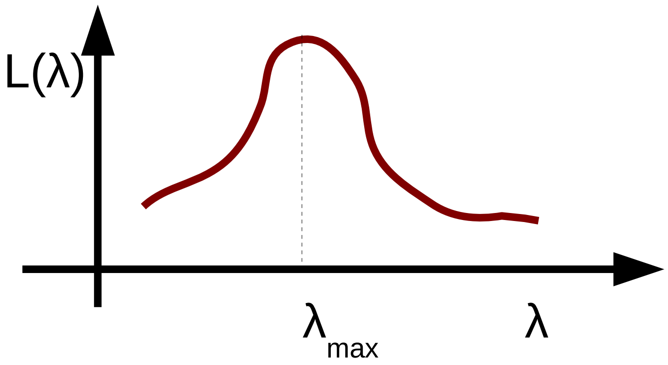
- **Funkcją wiarygodności** nazywamy iloczyn postaci:

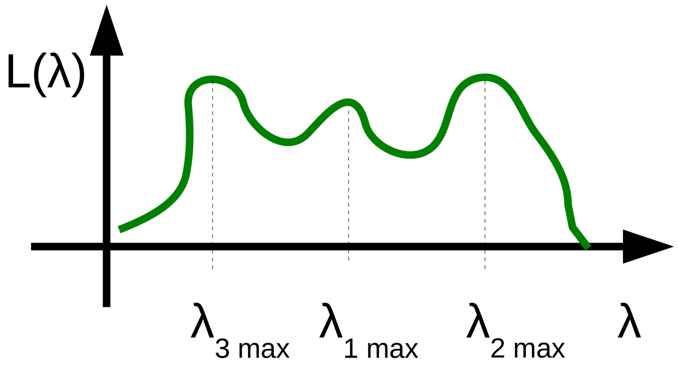
$$\boxed{L=\prod_{j=1}^{N} f\left(\boldsymbol{X}^{(j)};\boldsymbol{\lambda}\right)}$$

- **Jak wyznaczyć maksimum?**
  - warunek konieczny: przyrównać pierwszą pochodną L do zera
- Różniczkowanie iloczynu jest jednak niewygodne, wprowadzamy więc **logarytm funkcji wiarygodności** L:

$$l=\ln L=\sum_{j=1}^{N} \ln f\left(\boldsymbol{x}^{(j)};\boldsymbol{\lambda}\right)$$

- **Funkcją wiarygodności** jest odpowiednikiem gęstości prawdopodobieństwa, tylko określona dla parametrów. Ponieważ jest funkcją próby losowej, jest również zmienną losową





równania wiarygodności

$$\frac{\partial l}{\partial \lambda_i}=0, \quad i=1,2,\ldots,p$$

Przykład – estymacja rozkładu norm.

- **Przykład:** badamy rozkład wagi studentów
  - rozkład wagi studentek w populacji jest opisany rozkładem normalnym o wartości średniej μ i wariancji σ^2
  - załóżmy, że wybraliśmy N=10 reprezentantów, których wagi (w kg) $X^{(j)}$ układają się następująco:
    - 52,2 55,3 59,0 57,6 67,6 72,6 68,9 62,6 67,6 81,6
  - **Zadanie:** na podstawie wyniku pomiaru znajdź najbardziej wiarygodną estymację parametrów rozkładu
  - oczywiście funkcja rozkładu prawdopodobieństwa każdego wyniku (wagi studentek) dana jest wzorem:

$$f\left(x^{(j)};\mu,\sigma\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-\left(x^{(j)}-\mu\right)^2}{2\sigma^2}\right)$$

  - konstruujemy funkcję wiarygodności:

$$L(\hat{\mu},\hat{\sigma})=\sigma^{-n}(2\pi)^{-n/2}\exp\left[-\frac{1}{2\hat{\sigma}^2}\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)^2\right]$$

Przykład – estymacja rozkładu norm.

– oraz logarytmiczną funkcję wiarygodności:

$$l(\hat{\mu},\hat{\sigma}) = \ln L(\hat{\mu},\hat{\sigma}) = \ln\left(\hat{\sigma}^{-N}(2\pi)^{-n/2}\left[-\frac{1}{2\hat{\sigma}^2}\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)^2\right]\right) = -\frac{1}{2\hat{\sigma}^2}\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)^2 + const$$

– równania wiarygodności:

$$\boxed{\frac{dl(\hat{\mu},\hat{\sigma})}{d\hat{\mu}} = 0 \qquad \frac{dl(\hat{\mu},\hat{\sigma})}{d\hat{\sigma}} = 0}$$

$$\frac{dl(\hat{\mu},\hat{\sigma})}{d\hat{\mu}} = \frac{-2\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)(-1)}{2\sigma} = 0 \Rightarrow \sum_{j=1}^{N}\left(x^{(j)}\right) - N\hat{\mu} = 0 \Rightarrow \boxed{\hat{\mu} = \frac{\sum_{j=1}^{N} x^{(j)}}{n}}$$

$$\frac{dl(\hat{\mu},\hat{\sigma})}{d\hat{\sigma}} = -\frac{n}{2\hat{\sigma}} + \frac{\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)}{2\hat{\sigma}^2} = 0 \Rightarrow \boxed{\hat{\sigma}^2 = \frac{\sum_{j=1}^{N}\left(x^{(j)}-\hat{\mu}\right)^2}{n}}$$

padło mailowo pytanie, czemu nie n-1
- estymatorem największej wiarygodności jest **obciążony** estymator wariancji

– dla formalności powinniśmy jeszcze sprawdzić drugie pochodne...

Estymatory obciążone

- **Czasem nie da się wyznaczyć estymatora nieobciążonego:** jaki jest warunek osiągnięcia <u>minimalnej wariancji</u>?
- Można pokazać (Brandt), że zajdzie tak, gdy: $E(S)=B(\lambda)+\lambda$

$$\boxed{l'=A(\lambda)(S-E(S)) \qquad l=B(\lambda)S+C(\lambda)+D \qquad L=d\exp\left(B(\lambda)S+C(\lambda)\right)}$$

- Wtedy w przypadku <u>estymatora nieobciażonego</u> o minimalnej wariancji otrzymujemy:

$$\sigma^2(S)=\frac{1}{I(\lambda)}=\frac{1}{E(l'^2)}$$

$$\boxed{\sigma^2(S)=\frac{1}{|A(\lambda)|}}$$

- Estymatory spełniające powyższe warunki nazywamy **estymatorami o najniższej wariancji**
- Przykład – rozkład Poissona: $f(k)=\frac{\lambda^k}{k!}e^{-\lambda}$

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

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

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

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

Rozkład χ^2 – zastosowanie

- Rozkład χ^2 stosuje się jako **miarę ufności** uzyskanego wyniku (**odchylenia elementów próby od wartości średniej populacji)**. Im mniejsza wartość χ^2 tym pozornie słuszniejszy wynik. Jako miary zaufania do wyniku używa się wielkości: $W(\chi^2)=1-F(\chi^2)\equiv p=1-\alpha$

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

  nazywanej **poziomem ufności** (zwykle podawanym w % ilości odchyleń standardowych rozkładu normalnego σ)
  - wielkość α jest nazywana **poziomem istotności**

- W rzeczywistych przypadkach mamy do czynienia z pełnym rozkładem Gaussa o dowolnym a i σ. Wprowadzamy wtedy odpowiednie przeskalowanie

elementy próby losowej

$$\chi^2=X^2=\frac{(X_1-a)^2+(X_2-a)^2+\ldots+(X_n-a)^2}{\sigma^2}$$

- a w ogólnym przypadku wielowymiarowym, gdy zmienne są zależne:

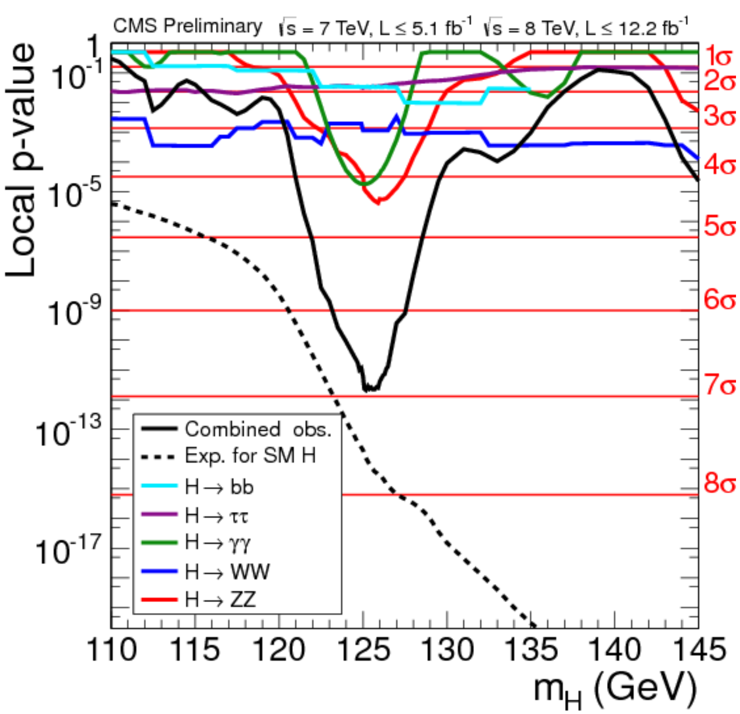
$$\chi^2=X^2=(\boldsymbol{X}-\boldsymbol{a})^T B(\boldsymbol{X}-\boldsymbol{a})$$

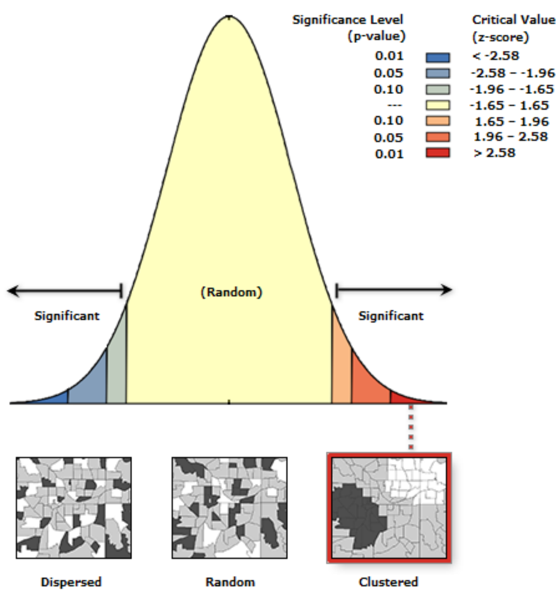
Rozkład χ^2 a estymator wariancji

- Można udowodnić, że zmienna losowa: $\frac{n-1}{\sigma^2} S^2$ estymatyor wariancji

  ma rozkład χ^2 z $f=n$-1 stopniami swobody. Wynika to stąd, że wyrażenia $(X_i - \bar{X})^2$ nie są liniowo niezależne (co już wiemy). Każde dodatkowe równanie (więzy) pomiędzy wyrażeniami $(X_i - \bar{X})^2$ redukuje liczbę stopni swobody o 1

- Wartości p– przykład:





$$P(|Y-a| \leqslant \sigma) = 68{,}3\% \quad P(|Y-a| > \sigma) = 31{,}7\%$$

$$P(|Y-a| \leqslant 2\sigma) = 95{,}4\% \quad P(|Y-a| > 2\sigma) = 4{,}6\%$$

$$P(|Y-a| \leqslant 3\sigma) = 99{,}8\% \quad P(|Y-a| > 3\sigma) = 0{,}2\%$$

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. $1/\sqrt{10}$

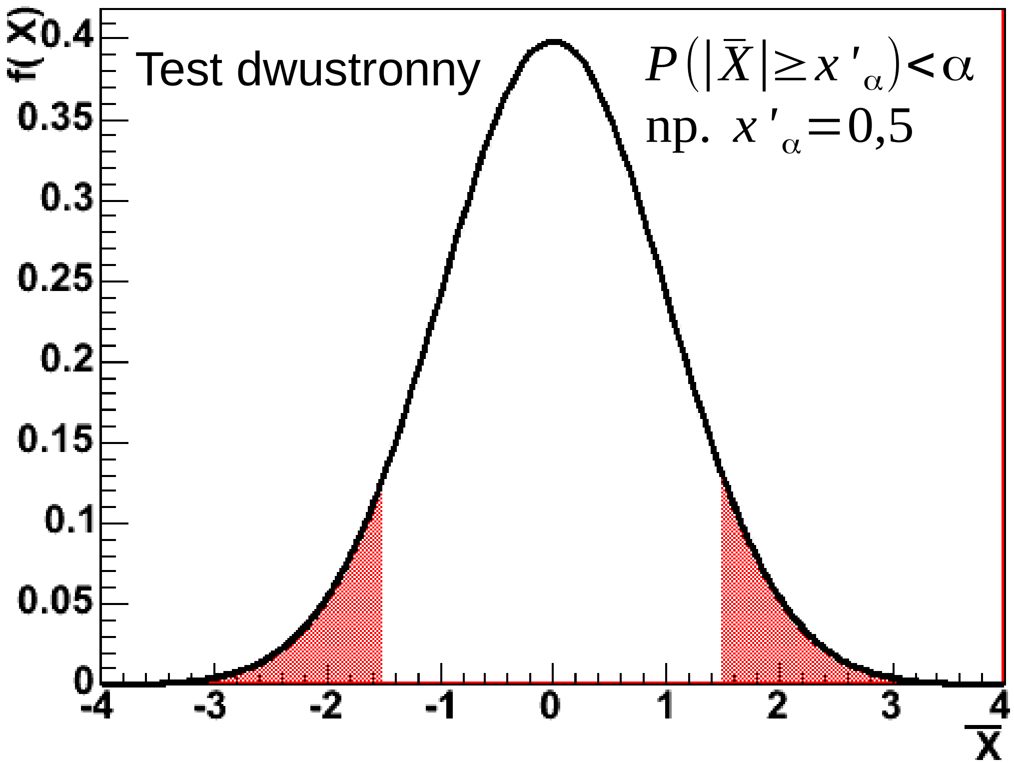
$$\sigma^2(\bar{X})=\frac{1}{n}\sigma^2(X)=\frac{1}{10}\cdot 1\Rightarrow\sqrt{\sigma^2(\bar{X})}=\frac{1}{\sqrt{10}}$$

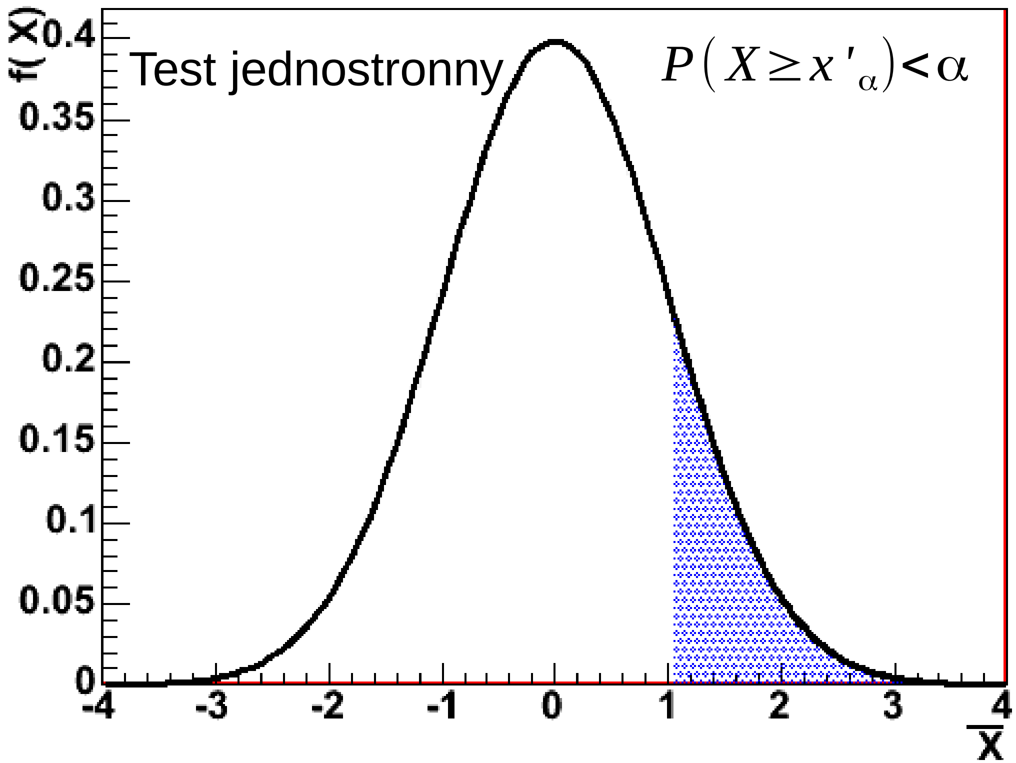
Weryfikacja hipotez statystycznych

- 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 spełniona** – jest mało prawdopodobne, aby próba pochodziła z rozkładu określonego przez testowaną hipotezę → **możemy ją odrzucić**
  - **prawdopodobieństwo zaobserwowania tego, że $|\bar{X}|$ jest duże, jest bardzo małe, ale takie nam się trafiło – więc prawdopodobnie (z prawdodobieństwem 1-α) 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 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:** $u_i = \dfrac{g_i - f_i}{\sigma_i}, \; i=1,2,\ldots,N$
- Jeśli nasza teoria (wartości f_i) jest prawdziwa, to <u>rozkłady różnic u_i mają postać standardowego rozkładu normalnego</u> **– 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: $\boxed{P(T > \chi^2_{1-\alpha}) = \alpha}$

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

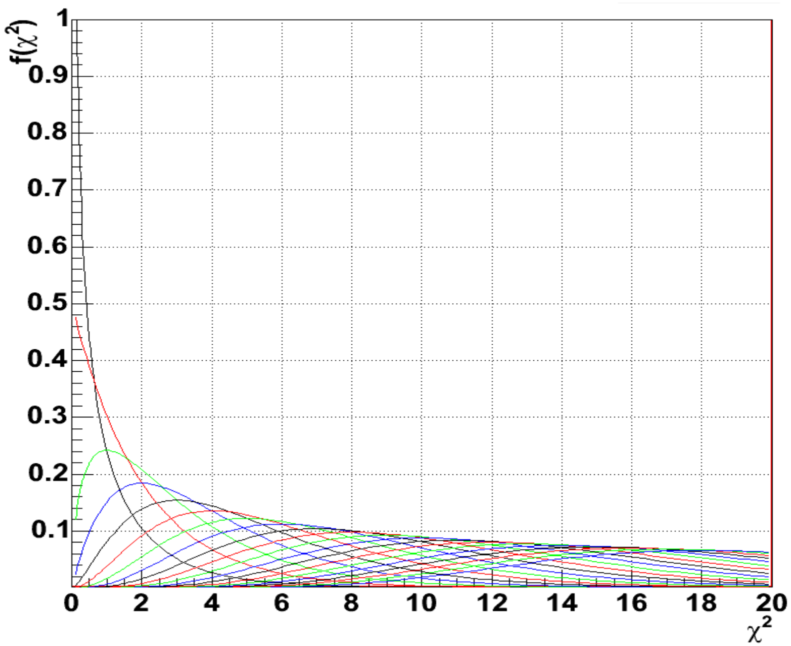
Test χ^2 dobroci dopasowania

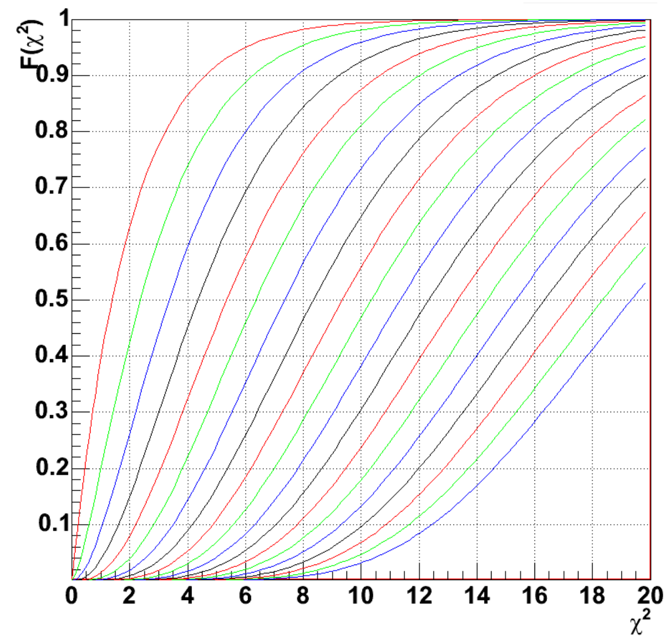
- **<u>Podsumowując</u>, w naszym przypadku musimy <u>dla danej realizacji próby (wyniku eksperymentu)</u> wyznaczyć wartość testową T i porównać ją z odpowiednim kwantylem rozkładu** χ^2 **o odpowiedniej liczbie stopni swobody:**

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

- **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 - 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.)

Test χ^2 - przykład

Zadanie

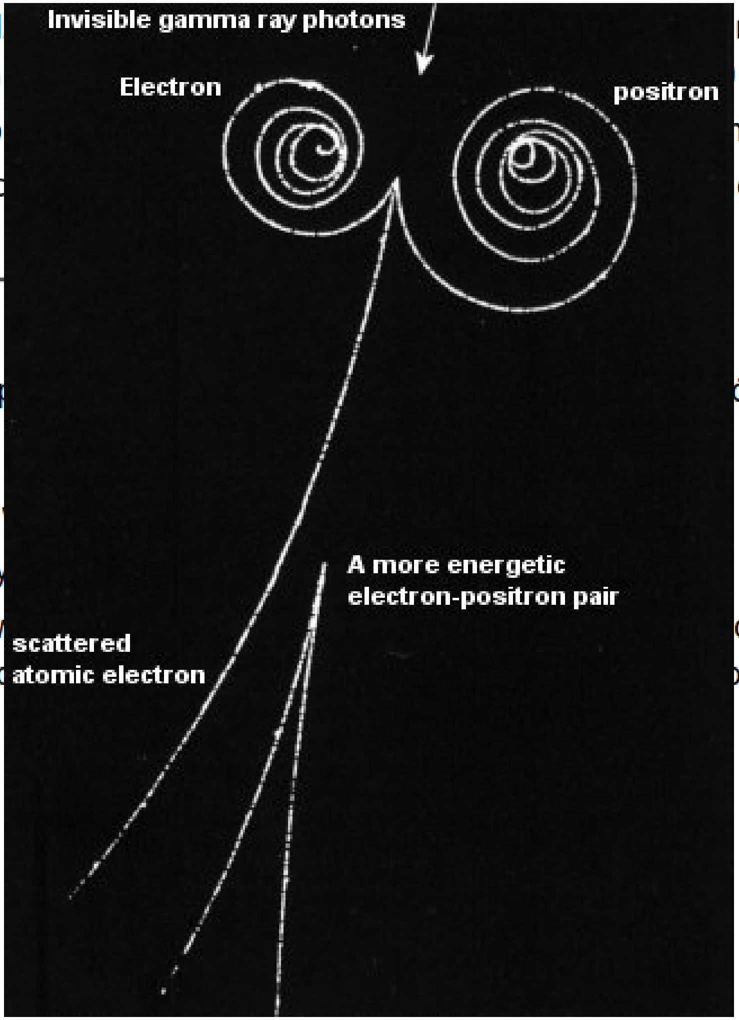
Weryfikacja hipotez statystycznych (5 pkt.)

- 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). 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 $T = \sum_k \frac{(n_k$

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

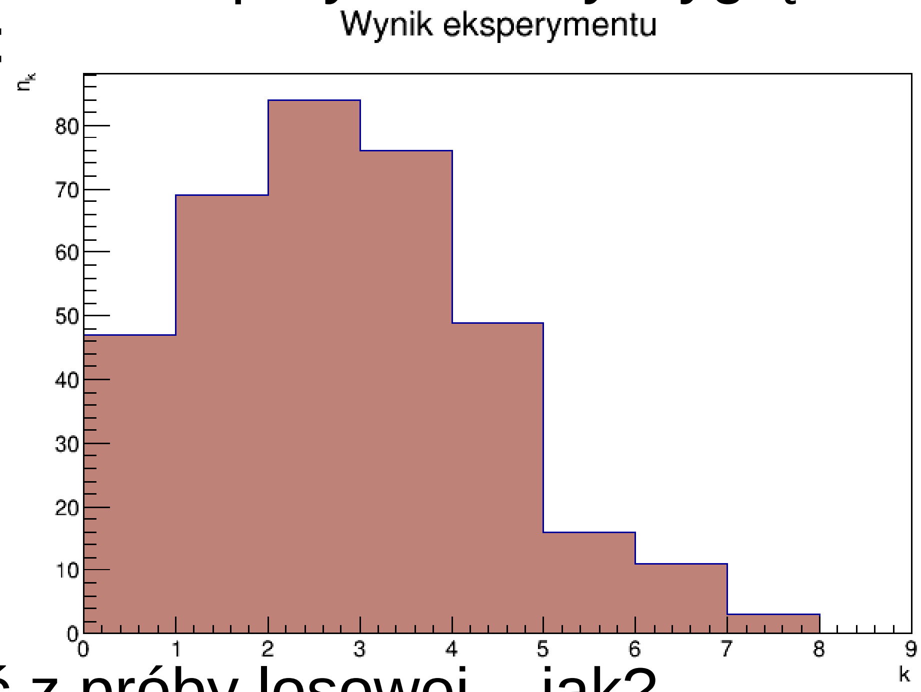
- Określić liczbę stopni swobody i obliczyć
- Zaimplementować funkcję zwracającą wy

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 (kilka slajdów temu):

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

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

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

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

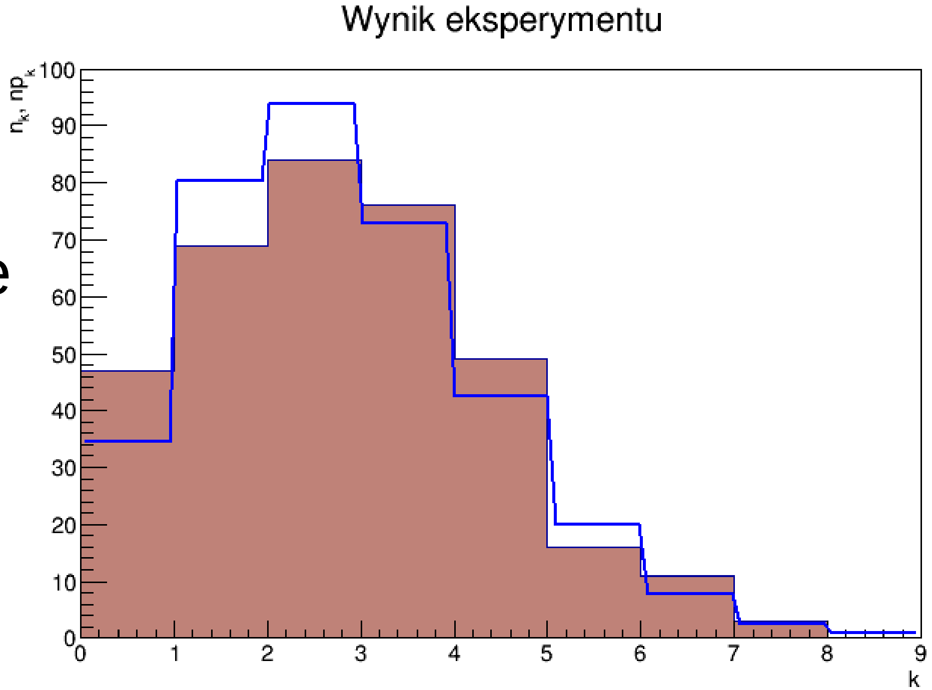
- Czyli estymatorem największej wiarygodności o minimalnej wariancji dla rozkładu Poissona jest średnia arytmetyczna z próby
- Oczywiście w naszym przypadku mamy histogram, który zawiera jakąś całkowitą liczbę wejść (całka z histogramu nie jest równa 1), wobec tego do średniej dodajemy wagi w postaci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

Hipotezy zerowa i alternatywna

Błędy I i II rodzaju

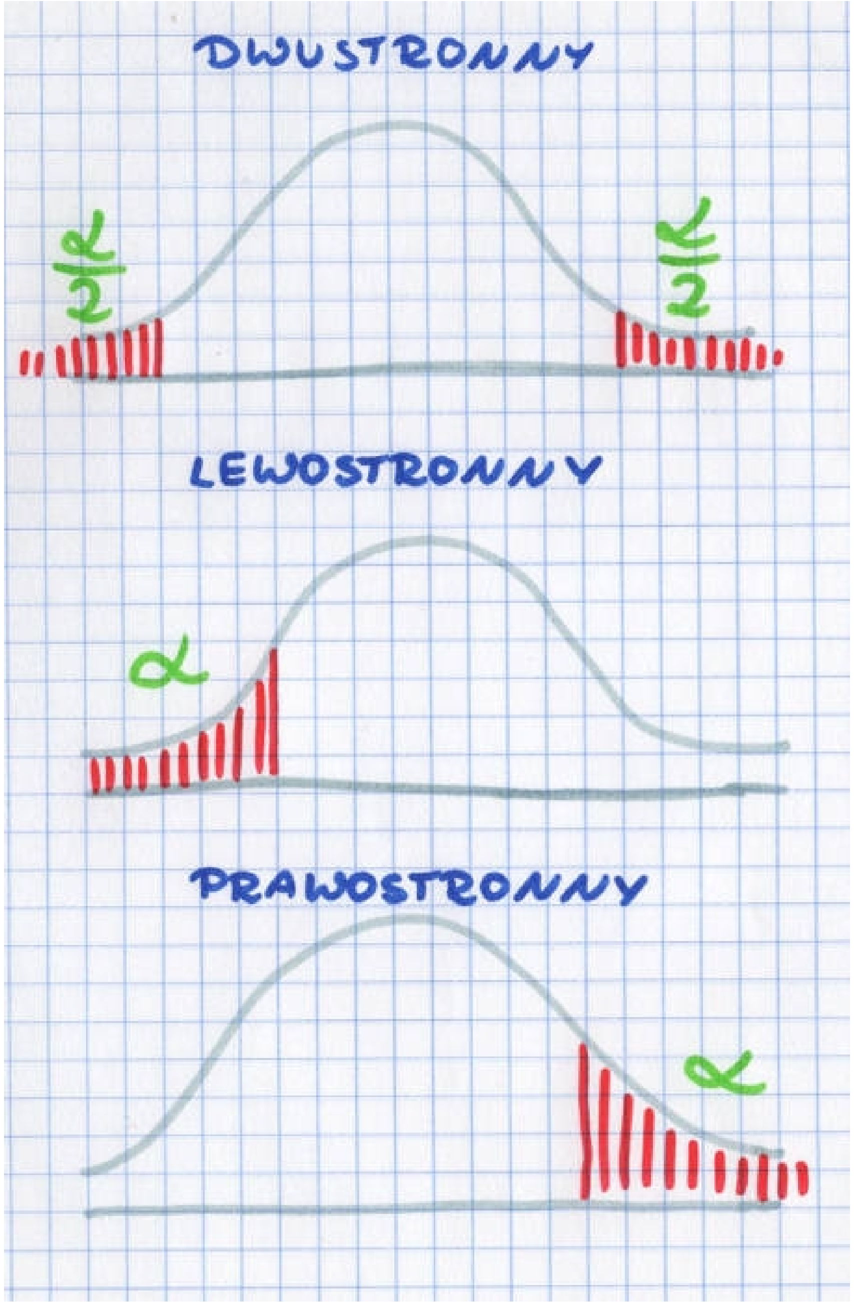
Na podstawie:
https://www.statystyczny.pl/hipotezy-statystyczne/

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 chi2 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 chi2 jest duże → teoria nie opisuje danych (test jednostronny – rozkład chi2 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

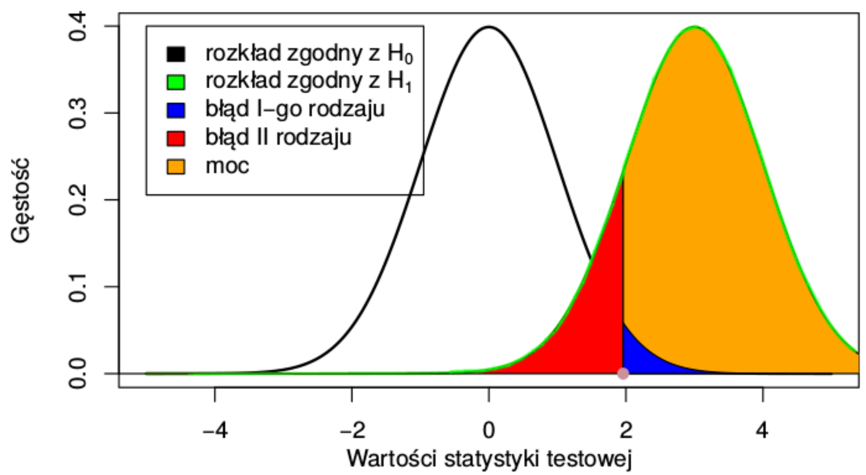
- **Błąd I rodzaju** to taki, gdzie odrzucamy hipotezę zerową a była ona prawdziwa
  - H_0: poziom szczęścia (średnio w populacji) po zjedzeniu dużej porcji lodów jest taki sam jak przed
  - H_A: poziom szczęścia (średnio w populacji) się zmienia
  - my na podstawie doświadczenia (próby losowej) zjedliśmy lody i np. przez te okropne wyrzuty sumienia, że znowu za dużo kalorii :) odrzucamy hipotezę zerową na rzecz alternatywnej
  - jeżeli w wyniku wielu prób losowych wynika, że jednak lądujemy w ogonie rozkładu owego szczęścia, 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

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

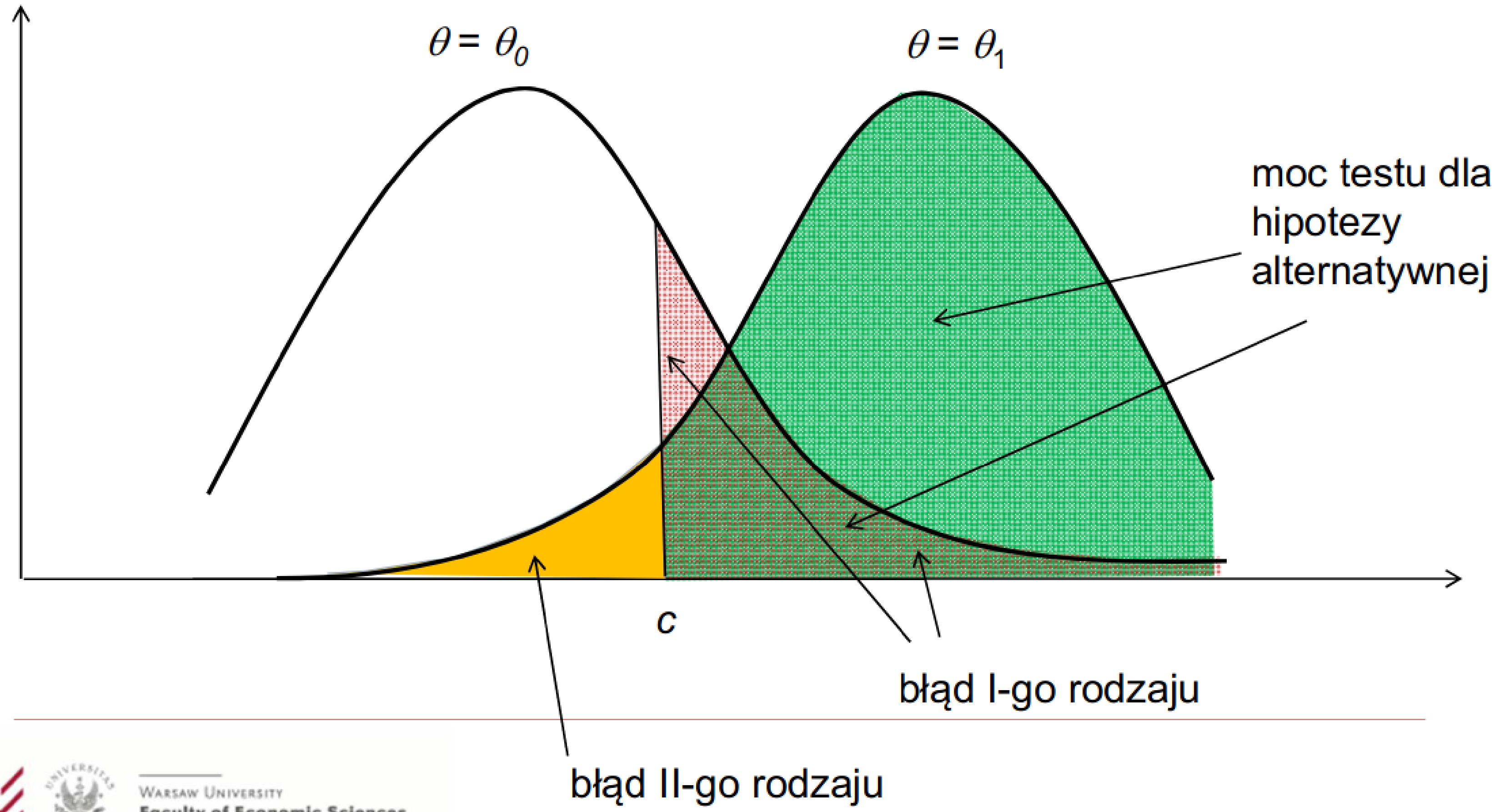
- Moc testu (prawdopodobieństwo, że prawidłowo odrzucimy hipotezę zerową) to 1-β. Inaczej mówiąc jest to prawdopodobieństwo niepopełnienia błędu II rodzaju.
- Moc testu zależy od kilku czynników:
  - Wielkości próby użytej w badaniu (im większa próba, tym większa moc testu).
  - Rzeczywistej wielkości 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).



Rys.2 Relacje między błędem I, II rodzaju oraz mocą

Moc testu:
interpretacja graficzna (1)

rozkłady statystyki testowej przy założeniu prawdziwości hipotezy zerowej i alternatywnej



Testy statystyczne

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

Przykładowe zadania na kolokwium z KADD (wykład):

1) Dana jest dystrybuanta zmiennej X:

x	(-inf, -2>	(-2, 1>	(1, 3>	(3, inf)
F(x)	0	0.2	0.8	1.0

Wyznacz funkcję prawdopodobieństwa.

2) Zmienna losowa X ma funkcję prawdopodobieństwa postaci:

x	-3	-1	3	5
p_i	0.1	0.2	0.5	0.2

Wyznaczyć funkcje prawdopodobieństwa zmienne U:
a) U = 2X +3
b) $U = X^3$
c) $U= X^2-5$
Ponadto wyznaczyć wartości oczekiwane i wariancję.

3) Zmienna losowa ma rozkład prawdopodobieństwa:
f(x) = c*sin(x) x in [0, π]
a) Dobrać stałą c, aby f(x) była gęstością prawdopodobieństwa
b) Wyznaczyć funkcję dystrybuanty
c) Wyznaczyć wartość oczekiwaną, wariancję, odchylenie standardowe
d) Wyznaczyć medianą

4) Zaproponować metodą do generacji liczb z rozkładu potęgowego

5) Zaproponować metodę do generacji liczb z rozkładu sinusoidalnego

6) Dwuwymiarowa zmienna losowa (X,Y) ma rozkład gęstości: $f(x,y) = 1/(20\pi) * \exp(-0.5*(x^2/4+y^2/25))$. Zbada c, czy zmienne X, Y są niezależne, Podać sposób na obliczenie (obliczyć): $P(-1<X<2, 0<Y<3)$

7) Dane są dwie próby:
   A) 21, 19, 14, 27, 25, 23,22, 18, 21 (N = 9)
   B) 16, 24,22, 21, 25, 21, 18 (N=7)

Czy przy poziomie istotności 5% wariancja próby (B) jest mniejsza niż próby (A)?

8) Zweryfikuj hipotezę, że 10 pomiarów zostało wylosowane z populacji o wartości średniej 25.6? Przyjmij, że poziom istotności wynosi 10%, załóż, że populacji ma rozkład normalny.
Pomiary: 22.03, 27.05, 27.5, 22.7, 25.3, 26.2, 27.9, 21.2, 23.3, 25.5.

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 $f \rightarrow \infty$
- 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**