Komputerowa Analiza Danych Doświadczalnych

Prowadząca:
dr inż. Hanna Zbroszczyk

e-mail: *gos@if.pw.edu.pl*
tel: +48 22 234 58 51
konsultacje: poniedziałek: 10-11, środa: 11-12

www: http://www.if.pw.edu.pl/~gos/students/kadd

Politechnika Warszawska
Wydział Fizyki
Pok. 117b (wejście przez 115)

METODA NAJWIĘKSZEJ WIARYGODNOŚCI

Metoda największej wiarygodności – Iloraz wiarygodności

Do tej pory: **estymacja wartości oczekiwanej, wariancji** (bez wnikania jak konstruować estymatory).

- p parametrów, $\boldsymbol{l}=(l_1, l_2, \ldots, l_p)$
- określone funkcją prawdopodobieństwa: $f=f(\boldsymbol{x}, \boldsymbol{l})$
- dla zmiennych losowych

Przeprowadzone **N doświadczeń** (lub pobrana próba N – elementowa),

pojedyncze doświadczenie (pomiar wielkości **x** lub pobranie próby o liczebności 1)

może doprowadzić do uzyskania wyniku $\mathbf{x}^{(j)}$.

Doświadczeniu przypisujemy liczbę: $\boldsymbol{x}=(x_1, x_2, \ldots, x_n)$

To **prawdopodobieństwo a posteriori**. Mówi ono po uzyskaniu wyniku, jakie było prawdopodobieństwo uzyskania tego właśnie wyniku (tzn. wartości $\mathbf{x}^{(j)}$).

$$dP^{(j)}=f(\boldsymbol{x}^{(j)}; \boldsymbol{l})\,dx$$

$$x_i^{(j)}<\boldsymbol{x}^{(j)}\leq x_i^{(j)}+dx_i^{(j)} \qquad i=1,2,\ldots,n$$

Metoda największej wiarygodności – Iloraz wiarygodności

Dla próby N (niezależnych) elementów, **prawdopodobieństwo uzyskania wyniku**:

$$\boldsymbol{x}^{(\boldsymbol{1})}, \boldsymbol{x}^{(\boldsymbol{2})}, \ldots, \boldsymbol{x}^{(\boldsymbol{N})}$$

Wynosi:

$$dP = \prod_{j=1}^{N} f(\boldsymbol{x}^{(\boldsymbol{j})}; \boldsymbol{l}) \, d\boldsymbol{x}$$

Gdy prawdopodobieństwo zależy od np. 2 parametrów:

$$Q = \frac{\prod_{j=1}^{N} f(\boldsymbol{x}^{(\boldsymbol{j})}; \boldsymbol{l}_{\boldsymbol{1}}) \, d\boldsymbol{x}}{\prod_{j=1}^{N} f(\boldsymbol{x}^{(\boldsymbol{j})}; \boldsymbol{l}_{\boldsymbol{2}}) \, d\boldsymbol{x}}$$

Zbiór parametrów typu pierwszego jest Q razy bardziej prawdopodobny

niż zbiór parametrów typu drugiego. Jest to **iloraz wiarygodności.**

$$L = \prod_{j=1}^{N} f(\boldsymbol{x}^{(\boldsymbol{j})}; \boldsymbol{l}) \, d\boldsymbol{x}$$

Funkcja wiarygodności - Jest zmienną losową.

Iloraz wiarygodności - przykład

Rzut asymetryczną monetą, na podstawie uzyskanych wyników próbujemy ustalić, czy moneta należy do klasy A, czy do klasy B monet.

	A	B
orzeł	1/3	2/3
reszka	2/3	1/3

Próba to 5 rzutów: 1 orzeł i 4 reszki.

$$L_A = \frac{1}{3}\left(\frac{2}{3}\right)^4$$

$$L_B = \frac{2}{3}\left(\frac{1}{3}\right)^4$$

$$Q = \frac{L_A}{L_B} = 8$$

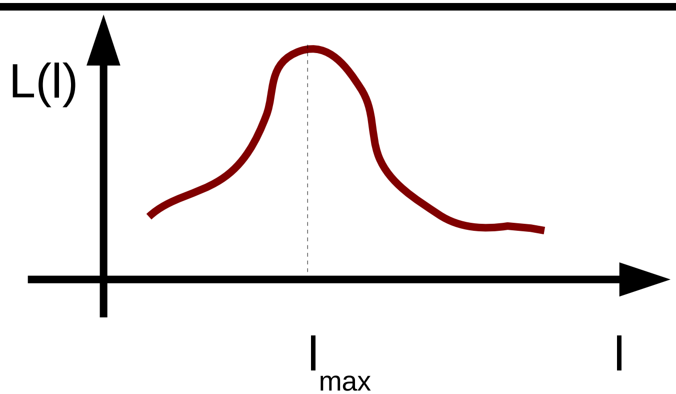
Z dużą dozą prawdopodobieństwa moneta należy do klasy monet A.

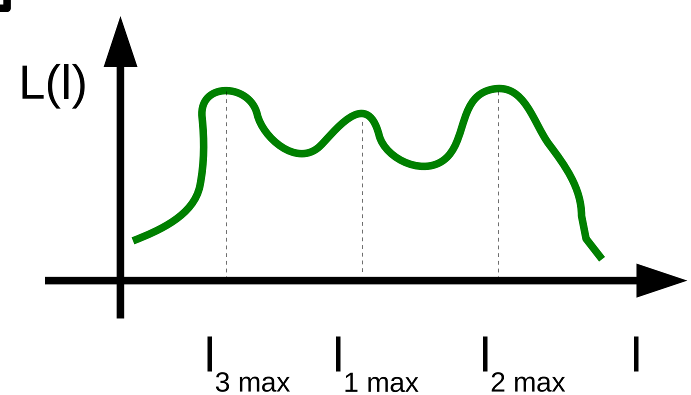
Metoda największej wiarygodności

- Największą **ufnością** obdarzony zbiór parametrów, którego funkcja wiarygodności osiąga maksymalną wartość.
- **Wyznaczenia maksimum**: przyrównać do zera pierwszą pochodną funkcji wiarygodności względem l_i.
- Ułatwienie obliczeń rachunkowych → **logarytm funkcji wiarygodności L**:

$$\ell = \ln L = \sum_{i=1}^{N} \ln f(x^{(j)}; \boldsymbol{l})$$

Logarytmiczna funkcja wiarygodności.





Metoda największej wiarygodności

Jeśli wektor **l** ma tylko jedną składową: l, należy rozwiązać **równanie wiarygodności**:

$$ł' = \frac{dł'}{dl} = 0$$

$$ł = \frac{\sum_{i=1}^{N} d \ln f(x^{(j)}; l)}{dl} = \frac{\sum_{i=1}^{N} f'}{f} = \sum_{i=1}^{N} \varphi(x^{(j)}; l)$$

$$\varphi(x^{(j)}; l) = \frac{\frac{d}{dl} f(x^{(j)}; l)}{f(x^{(j)}; l)}$$

W przypadku ogólnym, wektor **l** ma p składowych:

$$\frac{dł'}{dl_i} = 0 \qquad i = 1, 2, \ldots, p$$

Powtarzanie pomiarów o różnej dokładności - przykład

Pomiar danej wielkości różnymi przyrządami pomiarowymi

- pomiary będą układały się wokół **wartości rzeczywistej**,
- zakładamy, że **niepewności** będą miały **rozkład normalny**,
- **pojedynczy pomiar**: pobieranie próby o liczebności 1 z rozkładu normalnego o wart. średniej i war. σ^2.

Prawdopodobieństwo a posteriori dla uzyskanej wartości $x^{(j)}$:

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

Dla N pomiarów **funkcja wiarygodności**:

$$L=\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi}\,\sigma_j}\exp\left(-\frac{(x^{(j)}-l)^2}{2\sigma_j^2}\right)dx$$

Logarytm funkcji wiarygodności:

$$ł=-\frac{1}{2}\sum_{j=1}^{N}\frac{(x^{(j)}-l)^2}{\sigma_j^2}+const$$

Równanie wiarygodności:

$$\frac{dł}{dl}=\sum_{j=1}^{N}\frac{(x^{(j)}-l)^2}{\sigma_j^2}=0$$

$$l_{max}=\frac{\sum_{j=1}^{N}\frac{x^{(j)}}{\sigma_j^2}}{\sum_{j=1}^{N}\frac{1}{\sigma_j^2}}$$

Wynik najbardziej wiarygodny jest średnią ważoną N pomiarów, wagi są odwrotnościami wariancji.

Oszacowanie parametru N rozkładu hipergeometrycznego

Jeśli estymowany parametr może przyjmować wartości dyskretne:

Prawdopodobieństwo wyłowienia **n** ryb, z których **k** było zaznaczonych. **N** wszystkich ryb, **K** zaznaczone:

$$L(k;n,K,N)=\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Szukane N, dla którego L osiągnie maksimum.

Iloraz:

$$\frac{L(k;n,K,N)}{L(k;n,k,N-1)}=\frac{(N-n)((N-k))}{(N-n-K+k)N}$$

L osiąga **maksimum** dla wartości całkowitej, najbliższej **nK/k**

Nierówność informacyjna

Jak skonstruować estymator o optymalnych własnościach?

1) Estymator jest nieobciążony, jeśli **wartość obciążenia B(l) wynosi 0** dla każdej próby:

$$B(l)=E(S)-l=0$$

2) **Wariancja** estymatora ma być **jak najmniejsza**:

$$\sigma^2(S)\, minimalna$$

Często należy szukać **kompromisu** między minimalnym obciążeniem, a minimalną wariancją (ze względu na fakt istnienia nierówności informacyjnej).

Wariancja przybiera **minimalną** wartość, kiedy estymator **S** jest **stałą** (wariancja wynosi wtedy 0).

Jeśli estymator $S(x^{(1)}, x^{(2)}, \ldots, x^{(N)})$ jest funkcją próby $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$,to łączna **Gęstość prawdopodobieństwa** próby (losowej):

$$f(x^{(1)}, x^{(2)}, \ldots, x^{(N)}; l)=f(x^{(1)}; l) f(x^{(2)}; l) \ldots f(x^{(N)}; l)$$

Wartość oczekiwana:

$$E(S)=\int S(x^{(1)}, x^{(2)}, \ldots, x^{(N)}) f(x^{(1)}; l) f(x^{(2)}; l) \ldots f(x^{(N)}; l)\, dx^{(1)} dx^{(2)} \ldots dx^{(N)}$$

Nierówność informacyna, estymator o min. wariancji

$$E(S)=B(l)+l$$

Wyrażenie można różniczkować.

$$B'(l)+1$$

Można dowieść, że nierówność informacyjna:

$$\boxed{\frac{(B'(l)+1)^2}{I(l)} \leq \sigma^2(S)}$$

Informacja próby ze względu na l.

$$I(l)=E(ł'^2)=NE\{(\frac{f'(x;l)}{f(x;l)})^2\}$$

Nierówność informacyjna: związek między **obciążeniem**, **wariancją** i **informacją** zawartą w próbie.

Nie został wybrany żaden konkretny estymator, a nierówność stanowi ograniczenie od dołu dla wariancji estymatora. Jest to **ograniczenie dla minimalnej wariancji (ograniczenie Cramera-Rao)**.

Kiedy obciążenie nie zależy od l:

$$\sigma^2(S) \geq \frac{1}{I(l)}$$

Estymator o min. wariancji, estymator wystarczający

Jaki jest warunek **osiągnięcia minimalnej wariancji**, kiedy zachodzi wzór:

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

Można dowieść, że kiedy:

$$ł=A(l)(S-E(S))$$

A co za tym idzie:

$$l=\int ł\,dl=B(l)S+C(l)+D$$

$$L=d\exp(B(l)S+C(l)) \quad \textbf{(*)}$$

Estymatory związane z funkcja wiarygodności zapewniają minimalną wariancję (zgodnie

Z nierównością informacyjną) – ***estymatory o minimalnej wariancji*.**

Dla estymatora nieobciążonego, można dowieść, że:

$$\sigma^2(S)=\frac{1}{I(l)} \qquad \sigma^2(S)=\frac{1}{|A(l)|}$$

Jeśli zamiast relacji (*) zachodzi słabszy warunek: $L=g(S,l)c(x^{(1)},x^{(2)},\ldots,x^{(N)})$

To S jest **estymatorem wystarczającym** dla l.

Estymator parametru rozkładu Poissona

Rozkład Poissona:

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

Funkcja wiarygodności dla próby $k^{(1)}, k^{(2)}, \ldots, k^{(N)}$

$$ł=\sum_{j=1}^{N}[k^{(j)}\ln\lambda-\ln(k^{(j)}!)-\lambda]$$

$$\frac{dł'}{d\lambda}=ł'=\sum_{j=1}^{N}[\frac{k^{(j)}}{\lambda}-1]=\frac{1}{\lambda}\sum_{j=1}^{N}[k^{(j)}-\lambda]$$

$$ł=\frac{N}{\lambda}(\bar{k}-\lambda)$$

Średnia arytmetyczna k jest nieobciążonym estymatorem o minimalnej wariancji $\frac{\lambda}{N}$.

Estymator parametru rozkładu dwumianowego

Funkcja wiarygodności dla rozkładu dwumianowego o parametrach: p = λ oraz q = 1 – p = 1 - λ.

$$L(k;\lambda)=\frac{n!}{k!(n-k)!}\lambda^{k}(1-\lambda)^{n-k}$$

Otrzymujemy stąd:

$$ł=\ln L=k\ln\lambda+(n-k)\ln(1-\lambda)+\ln(\frac{n!}{k!(n-k)!})$$

$$ł'=\frac{k}{\lambda}-\frac{n-k}{1-\lambda}=\frac{n}{\lambda(1-\lambda)}(\frac{k}{n}-\lambda)$$

Średnia arytmetyczna k/n jest estymatorem o minimalnej wariancji $\lambda(1-\lambda)/n$

Prawo kombinacji błędów

Np. pomiar danej wielkości różnymi przyrządami pomiarowymi.

$$\frac{dł}{dl} = ł' = \sum_{j=1}^{N} \frac{x^{(j)} - l}{\sigma_j^2} \qquad ł' = \sum_{j=1}^{N} \frac{x^{(j)}}{\sigma_j^2} - \frac{\lambda}{\sigma_j^2} \qquad ł' = \sum_{j=1}^{N} \frac{1}{\sigma_j^2} \left\{ \frac{\sum \frac{x^{(j)}}{\sigma_j^2}}{\sum \frac{1}{\sigma_j^2}} - \lambda \right\}$$

$$S = \frac{\sum \frac{x^{(j)}}{\sigma_j^2}}{\sum \frac{1}{\sigma_j^2}}$$

To **nieobciążony estymator wartości średniej** l. Ma też minimalną wariancję.

Wariancja: $$\sigma(S) = \left(\sum_{j=1}^{N} \frac{1}{\sigma_j^2}\right)^{-1}$$

<u>Prawo kombinacji niepewności lub uśrednianie niepewności w kwadratach</u>.

$$\Delta(S) = \left(\frac{1}{(\Delta x_1)^2} + \frac{1}{(\Delta x_2)^2} + \ldots \frac{1}{(\Delta x_n)^2}\right)^{-1/2}$$

Kiedy jest jednakowa dokładność pomiarów:

$$S = \bar{x} \qquad \sigma^2(S) = \sigma^2 / n$$

Własności asymptotyczne estymatorów i f. wiarygodności

Próba jest duża (N→∞). **Estymator** S zdefiniowany jako **rozwiązanie równania wiarygodności**:

$$\frac{dł(l)}{dl} = ł'(l) = \sum_{j=1}^{N} \left(\frac{f'(x^{(j)};l)}{f(x^{(j)};l)} \right)_{l_{max}} = 0$$

Zakładając, że istnieje druga pochodna ł(l) względem l rozwijamy ł'(l) w **szereg Taylora**:

$$ł'(l) = ł'(l_{max}) + (l - l_{max}) ł''(l_{max}) + \dots$$

$$ł''(l_{max}) = \sum_{j=1}^{N} \left(\frac{f'(x^{(j)};l)}{f(x^{(j)};l)} \right)'_{l_{max}}$$

Wyrażenie to wartość średnia z próby; kiedy N duże zastąpić można **wartością oczekiwaną**:

$$ł''(l_{max}) = N\,E\left\{ \left(\frac{f'(x^{(j)};l)}{f(x^{(j)};l)} \right)'_{l_{max}} \right\}$$

$$ł''(l_{max}) = E(ł''(l_{max})) = -E(ł^2(l_{max})) = -I(l_{max}) = -1/b^2$$

b zależy jedynie od wartości **gęstości prawdopodobieństwa f i estymatora l_{max}**.

Własności asymptotyczne estymatorów i f. wiarygodności

Funkcję wiarygodności (k= const):

$$L(l)=k\exp\{-(l-l_{max})^2/2b^2\}$$

Można dowieść, że:

$$S=l_{max} \qquad E(S)=l$$

l_{max} To nieobciążony estymator o minimalnej wariancji:

$$\sigma^2(l_{max})=b^2=\frac{1}{I(l_{max})}=\frac{1}{E(ł'^2(l_{max}))}=-\frac{1}{E(ł''(l_{max}))}$$

Estymator ma ww. własności w granicznym przypadku dla dużych N.

Funkcja wiarygodności jest **asymptotycznie normalna.**

Funkcja wiarygodności L(l) to miara prawdopodobieństwa, że **wartość prawdziwa l_0 par. l = l.**

$$l=l_{max}\pm\sigma(l_{max})=l_{max}\pm\Delta l$$

Funkcja wiarygodności asymptotycznie dąży do rozkładu normalnego;

Prawdopodobieństwo tego, że **wartość prawdziwa**:

$$l_{max}-\Delta(l_{max})<l_0<l_{max}+\Delta l$$

wynosi 68.3%

Jednoczesna estymacja kilku parametrów

Mamy estymować jednocześnie p parametrów: $l = (l_1, l_2, \ldots, l_p)$. Zgodnie z równaniami:

$$\frac{\partial ł}{\partial l_i} = 0 \qquad i = 1, 2, \ldots, p$$

Otrzymamy **niepewności parametrów**, a nie same parametry.

Należy uwzględnić także korelacje między parametrami i ich niepewności.

logarytmiczna funkcja wiarygodności:

$$ł(x^{(1)}, x^{(2)}, \ldots, x^{(N)}; \boldsymbol{l}) = \sum_{i=1}^{N} \ln f(x^{(j)}; \boldsymbol{l})$$

rozwijając w **szereg Taylora** wokół punktu: $\boldsymbol{l}_{\boldsymbol{max}} = (l_{1\,max}, l_{2\,max}, \ldots, l_{p\,max})$

$$ł(\boldsymbol{l}) = ł(\boldsymbol{l}_{\boldsymbol{max}}) + \sum_{k=1}^{p} \left(\frac{\partial ł}{\partial l_k}\right)_{\boldsymbol{l}_{\boldsymbol{max}}} (l_k - l_{k\,max}) + \frac{1}{2} \sum_{n=1}^{p} \sum_{m=1}^{p} \left(\frac{\partial^2 ł}{\partial l_n \partial l_m}\right)_{\boldsymbol{l}_{\boldsymbol{max}}} (l_n - l_{max\,n})(l_m - l_{max\,m}) + \ldots$$

$$\left(\frac{\partial ł}{\partial l_k}\right)_{\boldsymbol{l}_{\boldsymbol{max}}} = 0 \qquad k = 1, 2, \ldots, p$$

Jednoczesna estymacja kilku parametrów

Rozwinięcie upraszcza się:

$$-(\ell(\boldsymbol{l})-\ell(\boldsymbol{l_{max}}))=\frac{1}{2}(\boldsymbol{l}-\boldsymbol{l_{max}})^T A(\boldsymbol{l}-\boldsymbol{l_{max}})+\dots$$

$$-A=\left.\begin{pmatrix} \frac{\partial^2 \ell}{\partial l_1^2} & \frac{\partial^2 \ell}{\partial l_1 \partial l_2} & \dots & \frac{\partial^2 \ell}{\partial l_1 \partial l_p} \\ \frac{\partial^2 \ell}{\partial l_1 \partial l_2} & \frac{\partial^2 \ell}{\partial l_2^2} & \dots & \frac{\partial^2 \ell}{\partial l_2 \partial l_p} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 \ell}{\partial l_1 \partial l_p} & \frac{\partial^2 \ell}{\partial l_2 \partial l_p} & \dots & \frac{\partial^2 \ell}{\partial l_p^2} \end{pmatrix}\right|_{l=l_{max}}$$

Jednoczesna estymacja kilku parametrów

Kiedy N→∞:

$$B=E(A)=-\left.\begin{pmatrix} E(\frac{\partial^2 ł}{\partial l_1^2}) & E(\frac{\partial^2 ł}{\partial l_1 \partial l_2}) & \dots & E(\frac{\partial^2 ł}{\partial l_1 \partial l_p}) \\ E(\frac{\partial^2 ł}{\partial l_1 \partial l_2}) & E(\frac{\partial^2 ł}{\partial l_2^2}) & \dots & E(\frac{\partial^2 ł}{\partial l_2 \partial l_p}) \\ \dots & \dots & \dots & \dots \\ E(\frac{\partial^2 ł}{\partial l_1 \partial l_p}) & E(\frac{\partial^2 ł}{\partial l_2 \partial l_p}) & \dots & E(\frac{\partial^2 ł}{\partial l_p^2}) \end{pmatrix}\right|_{l=l_{max}}$$

Funkcja wiarygodności: $L=k\exp\{-\frac{1}{2}(\boldsymbol{l}-\boldsymbol{l}_{max})^T(\boldsymbol{l}-\boldsymbol{l}_{max})\}$

Ma postać **p -wymiarowego rozkładu normalnego** z wartością średnią $\mathbf{l}_{\mathbf{max}}$ oraz z macierzą kowariancji

$C = B^{-1}$. Wariancje estymatorów to elementami diagonalne, a elementy pozadiagonalne opisują

kowariancje poszczególnych estymatorów.

Pierwiastek kwadratowy z wariancji to **odchylenie standardowe**.

Jednoczesna estymacja kilku parametrów

Za pomocą estymatora i jego błędu możliwe jest określenie **przedziału obejmującego wartość prawdziwą z prawdopodobieństwem 68.3%**. W przypadku wieloparametrowym funkcja wiarygodności to wielowymiarowy rozkład normalny, przedział ten określone jest nie tylko przez niepewności, ale poprzez pełną macierz kowariancji. Dla przypadku dwu-wymiarowego mamy do czynienia z elipsą kowariancji. Wykładnik rozkładu normalnego:

$$-\frac{1}{2}(\boldsymbol{l}-\boldsymbol{l}_{max})^T B(\boldsymbol{l}-\boldsymbol{l}_{max})=-\frac{1}{2}g(\boldsymbol{l})=-\{ł(\boldsymbol{l})-ł(\boldsymbol{l}_{max})\}$$

Elipsę kowariancji określa warunek:

$$g(\boldsymbol{l})=1=2\{ł(\boldsymbol{l})-ł(\boldsymbol{l}_{max})\}$$

Dla innych wartości g(l) otrzymuje się inne elipsoidy ufności.

Wewnątrz hiper - elipsoidy ufności – **obszar ufności**:

$$g(\boldsymbol{l})=1=2\{ł(\boldsymbol{l})-ł(\boldsymbol{l}_{max})\}=const$$

Wyznaczanie wartości średniej i wariancji r. normalnego

Celem jest **wyznaczenie wartości parametrów λ_1 (wartości średniej) oraz λ_2 (odch. standardowego)** z próby o liczebności N z rozkładu normalnego. (Pomiar zasięgu cząstek α w materii, który podlega rozkładowi normalnemu wokół wartości średniej).

Oba parametry są estymowane. **Funkcja wiarygodności:**

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

$$ł=-\frac{1}{2}\sum_{j=1}^{N}\frac{(x^{(j)}-\lambda_1)^2}{2\lambda_2^2}-N\ln\lambda_2-const$$

Układ równań wiarygodności:

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

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

Wyznaczanie wartości średniej i wariancji r. normalnego

Rozwiązania:

$$\lambda_1 = \frac{1}{N}\sum_{j=1}^{N} x^{(j)} \qquad \lambda_2 = \sqrt{\frac{\sum_{j=1}^{N}\left(x^{(j)}-\lambda_1\right)^2}{N}}$$

W celu wyznaczenia macierzy B:

$$\frac{\partial^2 ł}{\partial \lambda_1^2} = -\frac{N}{\lambda_2^2} \qquad \frac{\partial^2 ł}{\partial \lambda_1 \partial \lambda_2} = -\frac{2\sum x^{(j)}-\lambda_1}{\lambda_2^3} \qquad \frac{\partial^2 ł}{\partial \lambda_2^2} = -\frac{3\sum\left(x^{(j)}-\lambda_1\right)^2}{\lambda_2^4} + \frac{N}{\lambda_2^2}$$

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

Elementy diagonalne to **błędy** poszczególnych **parametrów**:

$$\Delta\lambda_1 = \lambda_2/\sqrt{N}$$

$$\Delta\lambda_2 = \lambda_2/\sqrt{2\mathrm{N}}$$

Oba parametry nie są ze sobą skorelowane.

WERYFIKACJA HIPOTEZ STATYSTYCZNYCH

Weryfikacja hipotez statystycznych

Poziom istotności α – określa prawdopodobieństwo popełnienia błędu I rodzaju

(błąd polegający na odrzuceniu hipotezy zerowej, która w rzeczywistości jest prawdziwa) ;

Maksymalny błąd, jaki jesteśmy w stanie zaakceptować ($\alpha = 0.01, 0.03, 0.05$)

Poziom ufności 1-α – określa prawdopodobieństwo wyznaczenia takiego przedziału ufności,

że rzeczywista wartość parametru znajduje się w tym przedziale ufności.

Weryfikacja hipotez statystycznych

- Wyznaczane są **wartość parametrów,** jakie nie są całkowicie nieznane.
- Są pewne **przypuszczenia**, co do ich wartości (przewidywania modeli, dotychczasowe wyniki).
- Zadanie próby (pomiaru) to **weryfikacja (test)** tej **hipotezy**.
- Procedura weryfikacji hipotez statystycznych to **test statystyczny.**

- Pobrano 10-elementową **próbę**, uzyskana średnia arytmetyczna $\bar{x}=0.154$
- **Założenie**: zmienna losowa pochodzi z populacji opisanej **rozkładem normalnym.**
- Jeśli przyjmiemy **hipotezę** za **słuszną**, zmienna losowa ma rozkład normalny z m= 0 oraz s $=1/\sqrt{10}$
- Jakie jest **prawdopodobieństwo** zaobserwowania wartości $|\bar{x}|\geq 0.154$?
- Korzystając z **tabel** dystrybuanty rozkładu normalnego:

$$P(|\bar{x}|\geq 0.154)=2\{1-\psi_0(0.154*\sqrt{10})\}=0.62$$

- Gdy hipoteza prawdziwa, istnieje prawdopodobieństwo 62%, że próba o liczebności 10 może mieć wartość średnią różniącą się więcej niż o 0.154 od wartości 0 średniej z populacji.
- **Hipoteza** jest **prawdziwa** czy nie? Odpowiedź musi być uzyskana w inny sposób.
- Źródłem trudności jest **probabilistyczny charakter wyników**, które można usunąć wprowadzając α.

Weryfikacja hipotez statystycznych

1) Ustalamy określoną wartość prawdopodobieństwa **α.**

2) Zakładając, że hipoteza prawdziwa pytamy:

prawdopodobieństwo, że zostaną zaobserwowane określone wartości próby < α ?

$$P(|\bar{x}| \geq 0.154) < \alpha$$

3) **Nierówność spełniona** → jest mało prawdopodobne, aby próba pochodziła z rozkładu określonego przez testowaną hipotezę statystyczną. Hipotezą odrzucamy.

4) **Nierówność niespełniona** → omawiane prawdopodobieństwo przewyższa wartość α, →

→ nie wnioskujemy, że hipoteza jest prawdziwa, lecz, że **nie ma podstaw do odrzucenia** hipotezy na zadanym poziomie istotności

(hipoteza nie jest sprzeczna z wynikiem uzyskanym wskutek poboru próby).

Wybór poziomu istotności α zależy od badanego problemu.

- Przy badaniu **kontroli jakości ołówków** można przyjąć 1%.
- W stawkach **ubezpieczeniowych** czasem wartość 0.01 jest za duża.
- W **analizie danych** zazwyczaj stosuje się wartości: 5%, 1%, 0,1%.

Weryfikacja hipotez statystycznych

- Na podstawie tablic statystycznych **dystrybuanty rozkładu normalnego** określamy granice dla dla $|\bar{x}|$ odpowiadające powyższym poziomom istotności:

$$0.05 = 2\{1-\psi_0(1.96)\} = 2\{1-\psi_0(0.62 * \sqrt{10})\}$$

$$0.01 = 2\{1-\psi_0(2.58)\} = 2\{1-\psi_0(0.82 * \sqrt{10})\}$$

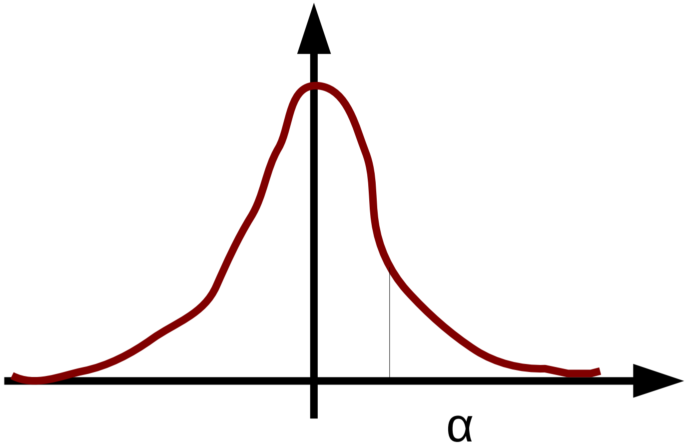
$$0.001 = 2\{1-\psi_0(3.29)\} = 2\{1-\psi_0(1.04 * \sqrt{10})\}$$

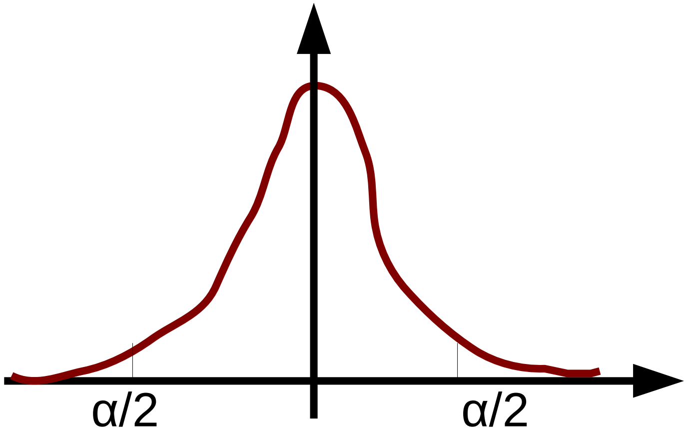
Do **odrzucenia hipotezy** wymagane jest, aby wartość średnia przekraczała odpowiednio: 0.62, 0.82, 1.04.

W niektórych przypadkach ważny jest **znak** rozpatrywanej wielkości wartości średniej.

Badamy **odchylenia** w jedną stronę i pytamy, czy spełniony jest warunek:

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





Test równości wariancji (test F-Fishera)

Porównywanie wariancji populacji o jednakowych wartościach średnich.

- **Problem**: pomiar tej samej wielkości dwoma przyrządami pomiarowymi.
- Jeśli pomiary będą miały **jednakowe wariancje** → **dokładność przyrządów taka sama.**
- **Założenie**: populacje mają rozkład normalny.
- Pobierane próby o **liczebności** N1 oraz N2.
- Dla obu liczona jest **wariancja** (estymator wariancji) i obliczany jest iloraz $F=\frac{s_1^2}{s_2^2}$
- Jeśli hipoteza o równości wariancji jest prawdziwa → F ~ 1.
- Dla każdej próby można skonstruować statystykę o rozkładzie X^2:

$$X_1^2=\frac{(N_1-1)s_1^2}{\sigma_1^2}=\frac{f_1 s_1^2}{\sigma_1^2} \qquad f_1=N_1-1$$

$$X_2^2=\frac{(N_2-1)s_1^2}{\sigma_2^2}=\frac{f_2 s_1^2}{\sigma_2^2} \qquad f_2=N_2-1$$

Liczby stopni swobody (ang. *NDF*)

Test równości wariancji (test F-Fishera)

Przy **równych** wariancjach:

$$F = \frac{f_2 X_1^2}{f_1 X_2^2}$$

Prawdopodobieństwo, że wartość ilorazu $\frac{X_1^2}{X_2^2}$ jest mniejsza od Q:

$$W(Q) = P\left(\frac{X_1^2}{X_2^2} < Q\right)$$

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

Można zdefiniować:

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

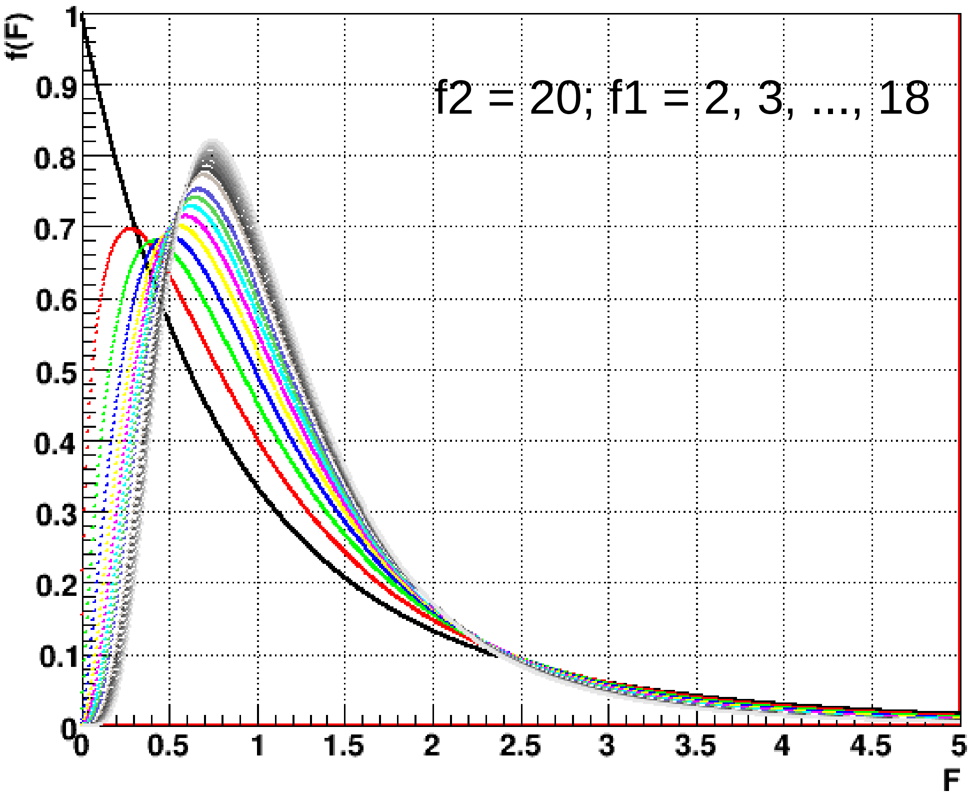
$$W(F) = P\left(\frac{s_1^2}{s_2^2} < F\right)$$

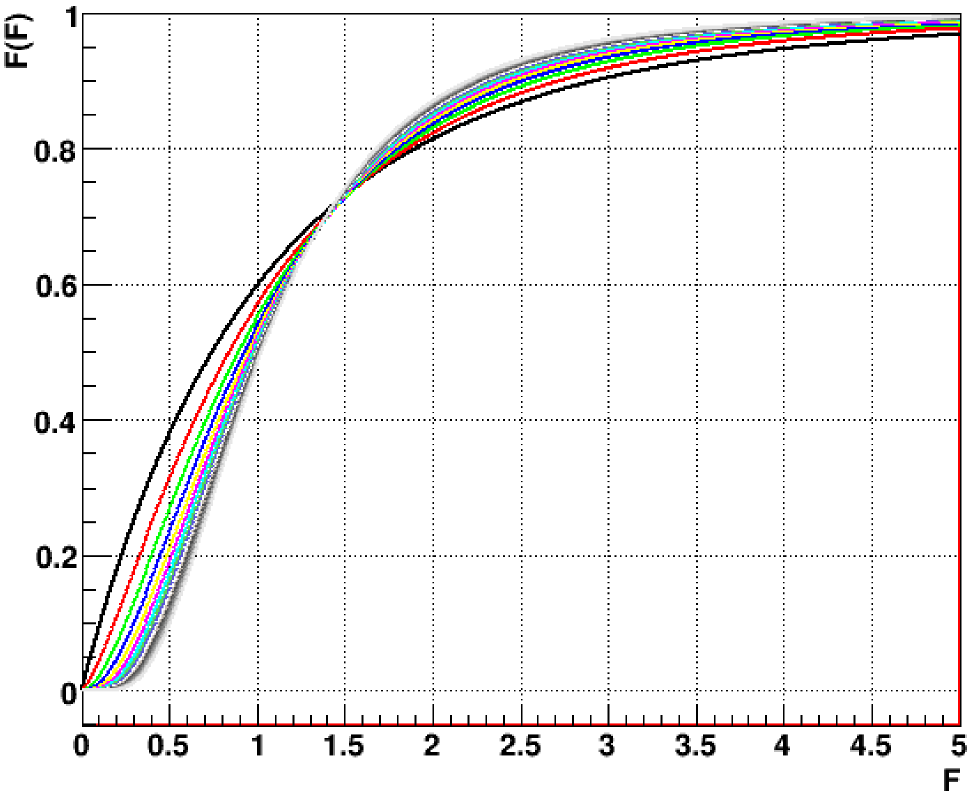
Dystrybuanta rozkładu Fishera

Test równości wariancji (test F-Fishera)

Gęstość prawdopodobieństwa:

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





Test równości wariancji (test F-Fishera)

Można **dowieść**, że dla $f_2 > 2$:

$$E(F) = \frac{f_2}{f_2 - 2}$$

Granica F'_α :

$$P\left(\frac{s_1^2}{s_2^2} > F'_\alpha\right) = \alpha$$

$F'_\alpha = F_{1-\alpha}$ to **kwantyl rozkładu F**.

$$P\left(\frac{s_1^2}{s_2^2} < F'_\alpha\right) = P\left(\frac{s_1^2}{s_2^2} < F_{1-\alpha}\right) = 1 - \alpha$$

- Jeśli przy weryfikacji hipotezy otrzymaliśmy wartość ilorazu wariancji **wyższą** niż niż $F'_\alpha = F_{1-\alpha}$

- **Hipoteza**: $\sigma_1^2 > \sigma_2^2$ prawdziwa dla poziomu istotności α.

- Wartości $F'_\alpha = F_{1-\alpha}$ są **stabelaryzowane**.

Test równości wariancji (test F-Fishera)

- Na ogół używa się **testu dwustronnego:**

iloraz wariancji znajduje się pomiędzy dwiema wartościami granicznymi F''_{α} oraz F'''_{α} :

$$P\left(\frac{s_1^2}{s_2^2} > F''_{\alpha}\right) = \frac{1}{2}\alpha \qquad P\left(\frac{s_1^2}{s_2^2} < F'''_{\alpha}\right) = \frac{1}{2}\alpha$$

- Ponieważ F jest ilorazem to:

$$\frac{s_1^2}{s_2^2} < F'''_{\alpha}(f_1, f_2) \qquad \frac{s_2^2}{s_1^2} > F'''_{\alpha}(f_2, f_1)$$

- Pierwszy argument odpowiada: ilości stopni swobody dla rozkładu drugiego,

drugi argument: ilości stopni swobody dla rozkładu pierwszego.

$$P\left(\frac{s_1^2}{s_2^2} > F''_{\alpha}(f_1, f_2)\right) = \frac{1}{2}\alpha \qquad P\left(\frac{s_2^2}{s_1^2} > F''_{\alpha}(f_2, f_1)\right) = \frac{1}{2}\alpha$$

- Z wartości tablicowych → $F_{1 - ½\alpha}$ >1 dla wszystkich rozsądnych wielkości α.

- Weryfikować należy tylko **hipotezę**:

$$\frac{s_g^2}{s_k^2} > f_{1-\alpha/2}(f_g, f_k) \qquad s_g^2 > s_k^2$$

- Nierówność spełniona → **hipotezę odrzucamy**.

Test równości wariancji (test F-Fishera) - przykład

Za pomocą dwóch przyrządów zmierzono odległość (100cm):

Numer pomiaru	Przyrząd 1	Przyrząd 2
1	100	97
2	101	102
3	103	103
4	98	96
5	97	100
6	98	101
7	102	100
8	101	
9	99	
10	101	
Wartość średnia	100	99.8
NDF	9	6
Wariancja	34/9 = 3.7	39/6 = 6.5

F= 6.5/3.7 = 1.8

$$F''_{0.1}(6,9) = F_{0.95}(6,9) = 3.37 \qquad 1.8 < 3.37$$

Nie ma podstaw do odrzucenia hipotezy

Test Studenta. Porównanie wartości średnich

- **Zmienna losowa o rozkładzie normalnym**.
- Pobierana próba o liczebności **N** oraz znanej **wartości średniej.**
- **Wariancja zmiennej losowej:**

$$\sigma^2(\bar{x}) = \frac{\sigma^2(x)}{N}$$

- Dla dostatecznie dużych prób: wartość średnia z próby (na mocy centralnego twierdzenia granicznego) ma **rozkład normalny** ($\hat{x}$, $\sigma(\bar{x})$).
- Zmienna: $y = \frac{\bar{x} - \hat{x}}{\sigma(\bar{x})}$ opisuje standardowy rozkład Gaussa.

- Na ogół wartość **odchylenia standardowego** nieznana. **Estymator wariancji**:

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

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

- **Pytanie**: Jak bardzo wielkość znormalizowana będzie odbiegać od standardowego rozkładu normalnego, jeśli odchylenie standardowe zastąpić pierwiastkiem kwadratowym z estymatora wariancji?

  Jako, że zawsze można przesunąć początek układu współrzędnych, to możliwe jest zawsze uzyskanie $\hat{x} = 0$

Test Studenta. Porównanie wartości średnich

Rozważmy **zmienną**:

$$t = \frac{\bar{x}}{s_{\bar{x}}} = \frac{\bar{x}\sqrt{N}}{s_x}$$

Ponieważ **wielkość** $(N-1)s_x^2 = f\, s_x^2$ ma rozkład X^2 o licznie stopni swobody f= N-1:

$$t = \frac{\bar{x}\sqrt{N}\sqrt{f}}{X}$$

$$F(t) = P(t < t) = P\left(\frac{\bar{x}\sqrt{N}\sqrt{f}}{X} < t\right)$$

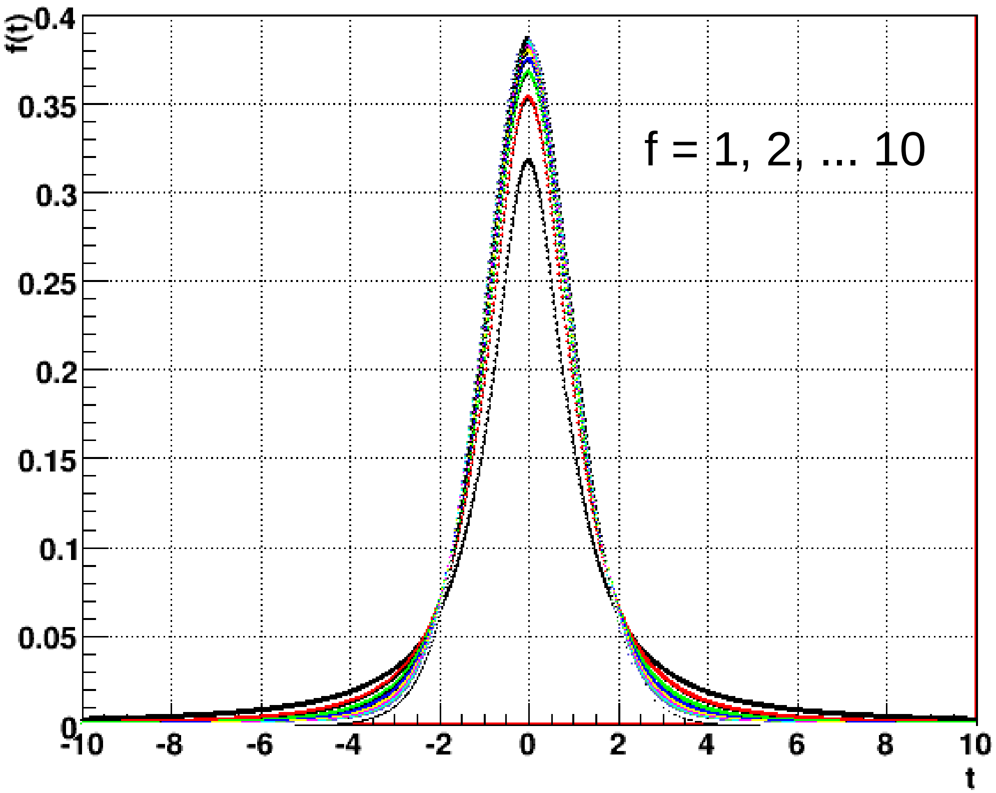
Można **dowieść**, że:

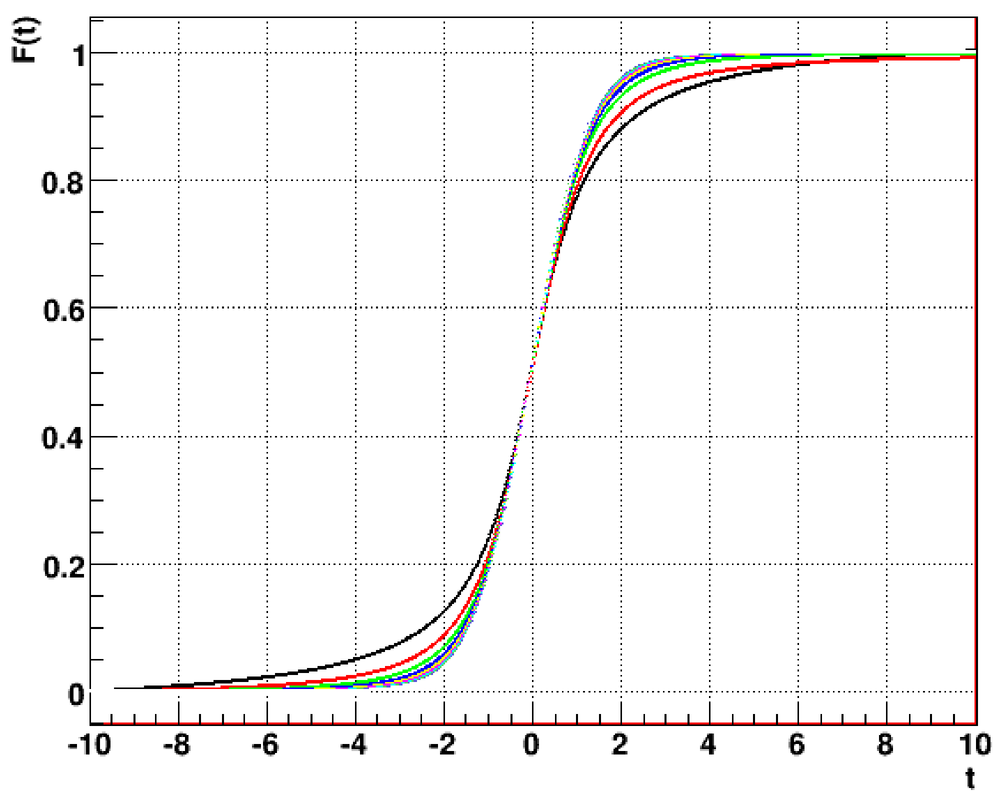
$$F(t) = \frac{\Gamma\left(\frac{1}{2}(f+1)\right)}{\Gamma\left(\frac{1}{2}f\right)\sqrt{\pi}\sqrt{f}} \int_{-\infty}^{t} \left(1+\frac{t^2}{f}\right)^{-\frac{1}{2}(f+1)} dt$$

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

Test Studenta. Porównanie wartości średnich

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





Test Studenta. Porównanie wartości średnich

- Rozważmy **zmienną**:

$$P(|t| \leq t) = 2F(|t|) - 1$$

- **Wartości graniczne** odpowiadające danemu poziomowi istotności:

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

- **Kwantyle** są stabelaryzowane.

ZASTOSOWANIE:

- **Hipoteza**: wartość oczekiwana z populacji mającej rozkład normalny równa jest L.
- Pobieramy próbę o liczebności N, wartości oczekiwanej oraz wariancji (policzonych).
- Jeśli przy zadanym poziomie istotności zachodzi warunek:

$$|t| = \frac{|\bar{x} - L|\sqrt{N}}{s_x} > t'_\alpha = t_{1-\frac{1}{2}\alpha}$$

- To hipotezę **odrzucamy** (test dwustronny).

$$|t| = \frac{|\bar{x} \pm L|\sqrt{N}}{s_x} > t'_{2\alpha} = t_{1-\alpha}$$ (test jednostronny).

Uogólnienie testu Studenta (porównanie w. średnich)

- Pobieramy próby o **liczebności** N1 oraz N2 z **2 populacj**i X oraz Y.

- **Hipoteza**: **równość wartości średnich**: $\hat{x} = \hat{y}$

- Z **CTG** → że wartości średnie mają rozkład normalny.

- **Wariancje**: $\sigma^2(\bar{x}) = \frac{1}{N_1}\sigma^2(x)$ $\quad \sigma^2(\bar{y}) = \frac{1}{N_1}\sigma^2(y)$

- **Estymatory wariancji**:

$$s_{\bar{x}}^2 = \frac{1}{N_1(N_1-1)}\sum_{j=1}^{N_1}(x_j-\bar{x})^2 \qquad s_{\bar{y}}^2 = \frac{1}{N_2(N_2-1)}\sum_{j=1}^{N_2}(y_j-\bar{y})^2$$

$\Delta = \bar{x} - \bar{y}$ Ma rozkład zbliżony do rozkładu normalnego.

$$\sigma^2(\Delta) = \sigma^2(\bar{x}) + \sigma^2(\bar{y})$$

- Jeśli hipoteza jest prawdziwa to $\bar{\Delta} = 0$. Podlega rozkładowi normalnemu: $\frac{\bar{\Delta}}{\sigma(\Delta)}$

- **Hipoteza** zakłada (na ogół), że **wartości średnie** zostały uzyskane z tej samej populacji,
czyli obie wariancje są identyczne.

Uogólnienie testu Studenta (porównanie w. średnich)

- **Estymator**:

$$s^2 = \frac{(N_1 - 1)\,s_x^2 + (N_2 - 1)\,s_y^2}{(N_1 - 1) + (N_2 - 1)}$$

$$s_{\bar{x}} = \frac{s^2}{N_1} \qquad s_{\bar{y}} = \frac{s^2}{N_2} \qquad s_\Delta^2 = s_{\bar{x}}^2 + s_{\bar{y}}^2 = \frac{N_1 + N_2}{N_1 N_2}\, s^2$$

- Można udowodnić, że:

Δ / s_Δ Podlega **rozkładowi Studenta** z liczbą stopni swobody: f = N1 + N2 -2.

- Równość wartości średnich można zweryfikować tzw. **testem różnic Studenta.**

- Powyższa wielkość liczona na podstawie wyników dwóch prób. Jej wartość bezwględną porównujemy z kwantylem rozkładu Studenta o liczbie stopni swobody f = N1 + N2 -2 dla zadanego poziomu istotności α. Jeśli zachodzi **nierówność**:

$$|t| = \frac{|\Delta|}{s_\Delta} = \frac{\bar{x} - \bar{y}}{s_\Delta} \geq t'_\alpha = t_{1 - \frac{1}{2}\alpha}$$

- **Hipotezę** o równości wartości średnich należy **odrzucić**.

Test równości wartości średnich (test Studenta) - przykład

Wyniki pomiarów stężenia kwasu neuraminowego w czerwonych ciałkach w krwi pacjentów zmarłych wskutek pewnej choroby (kolumna A) i dla grupy porównawczej pacjentów zdrowych (kolumna B).

A	B
21	16
24	20
18	22
19	19
25	18
17	19
18	19
22	
21	
23	
18	
13	
16	
23	
22	
24	
N1 = 16	N2 = 7
w. średnia = 20.3	w. średnia = 19.0
Wariancja = 171.8/15	Wariancja = 20/6

$$|\Delta| = |\bar{x} - \bar{y}| = 1.3$$

$$s^2 = \frac{15 s_x^2 + 6 s_y^2}{21} = 9.15$$

$$s_\Delta^2 = \frac{23}{112} s^2 = 1.88$$

$$\alpha = 5$$

$$f = 21$$

$$t_{1-\alpha/2} = 2.08$$

$$1.88 < 2.08$$

Materiał doświadczalny
Nie jest wystarczający.

Test X^2 z maksymalna liczbą stopni swobody

- **N** pomiarów g_i: i= 1,2, ..., N wraz z błędami σ_i.

- Wartość g_i jest **wynikiem pomiaru prawdziwej**, nieznanej wielkości h_i.

$$\mathbf{h_i = g_i + \varepsilon_i.}$$

- ε_i to **wielkość losowa** o rozkładzie normalnym o $m_i = 0$ oraz σ_i.

- Weryfikowana hipoteza: określa wartości h_i. Zakładamy, że $h_i = f_i$ oraz i=1, 2 ,..,N.

- Jeśli hipoteza jest prawdziwa, to wszystkie wielkości $u_i = \frac{g_i - f_i}{\sigma_i} \quad i = 1, 2, \ldots, N$ będą miały rozkład Gaussa.

- **Wielkość**:

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

Podlega rozkładowi X^2 o N stopniach swobody.

- Hipoteza fałszywa → poszczególne wartości u_i przyjmują duże wartości.

- Hipoteza odrzucona → dla zadanego poziomu istotności α spełniona jest zależność: $T > X^2_{1-\alpha}$

$X^2_{1-\alpha}$ - kwantyl rozkładu X^2 o N stopniach swobody.

Test X^2 z mniejszą liczbą stopni swobody

- Załóżmy, że wielkość g (przedmiot pomiaru) tot **funkcja niezależnej zmiennej kontrolowanej** t, (której wartość znamy g = g(t)).

- Poszczególne pomiary g_i odpowiadają pewnym wartościom t_i.

  $\mathbf{h_i. = h(t_i)}$.

- **Najprostszy** przypadek hipotezy, kiedy: f(t) = h(t) = at + b.

- **Hipoteza** może określać wartości liczbowe parametrów a, b.

- Znamy wtedy wszystkie **wartości** f_i, kiedy $f_i = h_i$.

- Hipoteza **prawdziwa** → wielkość T podlega rozkładowi X_2 o N stopniach swobody.

- Hipoteza zakłada istnienie **związku liniowego**, pomiędzy zmienną kontrolowaną, a zmienną h → → parametry a oraz b są nieznane. Budowane są ich **estymatory**: (będące funkcjami zmierzonych wartości g_i oraz ich błędów σ_i).

- **Hipoteza**: $h_i = h(t_i) = f_i = \tilde{a}\, t_i + \tilde{b}$

- **Estymatory** (funkcje zmierzonych wartości), nie zawsze są unormowane oraz niezależne.
- W celu wyznaczenia wartości estymatorów wprowadza się 2 **równania więzów**,
- **NDF** ulega zmniejszeniu do N-2.

Test X^2 i doświadczalny rozkład gęstości

- Zakładamy istnienie **dystrybuanty** F(x) odpowiadającej gęstości prawdopodobieństwa f(x).

- Zakres zmienności zmiennej losowej *x* dzielimy na r **przedziałów**: $\xi_1, \xi_2, \ldots, \xi_r$

- Poprzez całkowanie f(x) w poszczególnych przedziałach dostajemy **prawdopodobieństwo** zaobserwowania zmiennej w danym przedziale: $p_i = P(x \in \xi_i) = \int_{\xi_i} f(x)\,dx \qquad \sum_{i=1}^{r} p_i = 1$

- Z pobranej próby o liczebności n oznaczamy przez n_i liczbę takich elementów próby, jakie znalazły się w danym przedziale. Zachodzi relacja: $\sum_{i=1}^{r} n_i = n$

- Na podstawie hipotezy dot. gęstości prawdopodobieństwa: $n_i = n\,p_i$.

- Przedmiot weryfikacji: **hipoteza** zakładająca, że dla dużych wartości liczb n_i, ich wariancja wynosi n_i oraz, że rozkład u_i można przybliżyć rozkładem Gaussa. $u_i^2 = \frac{(n_i - np_i)^2}{n_i}$

- Słuszne jest: $$u_i^2 = \frac{(n_i - np_i)^2}{n\,p_i}$$

- **Suma kwadratów**: $$X^2 = \sum_{i=1}^{r} u_i^2 = \frac{\sum_{i=1}^{r} (n_i - np_i)^2}{n\,p_i}$$

- Jeśli przyjęta hipoteza jest prawdziwa, to, że statystyka X^2 ma (dla dużych n) rozkład X^2 .

- Zmienne nie są niezależne, **NDF** = r-1.

- Jeśli doświadczenie pozwala wyznaczyć p parametrów, to **NDF** zmniejsza się do f = r - 1 - p.

Test X^2 i dopasowanie rozkładu Poissona

W oddziaływaniach **fotonów z protonami**, wodorowa komora pęcherzykowa naświetlona została wiązką fotonów o wielkiej energii. Fotony konwertują w pary e+e-. Częstość występowania zdjęć z 0, 1, 2, parami e+e- powinna podlegać rozkładowi Poissona.

Estymator o największej wiarygodności parametru z rozkładu Poissona:

$$\tilde{\lambda} = \frac{\sum_k k\, n_k}{\sum_k n_k} = 2.33$$

Liczba par elektronowych na zdjęciu k	Liczba zdjęć zawierających k par	Przewidywania Rozkładu Poissona $n\,p_k$	$\frac{(n_k - n\,p_k)^2}{n\,p_k}$
0	47	34.4	4.61
1	69	80.2	1.56
2	84	93.7	1.00
3	76	72.8	0.14
4	49	42.6	0.96
5	16	19.9	0.76
6	11	7.8	1.31
7	3	2.5	0.10
8	-	(0.7)	-
	n = 355		X^2 = 10.44

Liczba stopni swobody:

N= r – 1

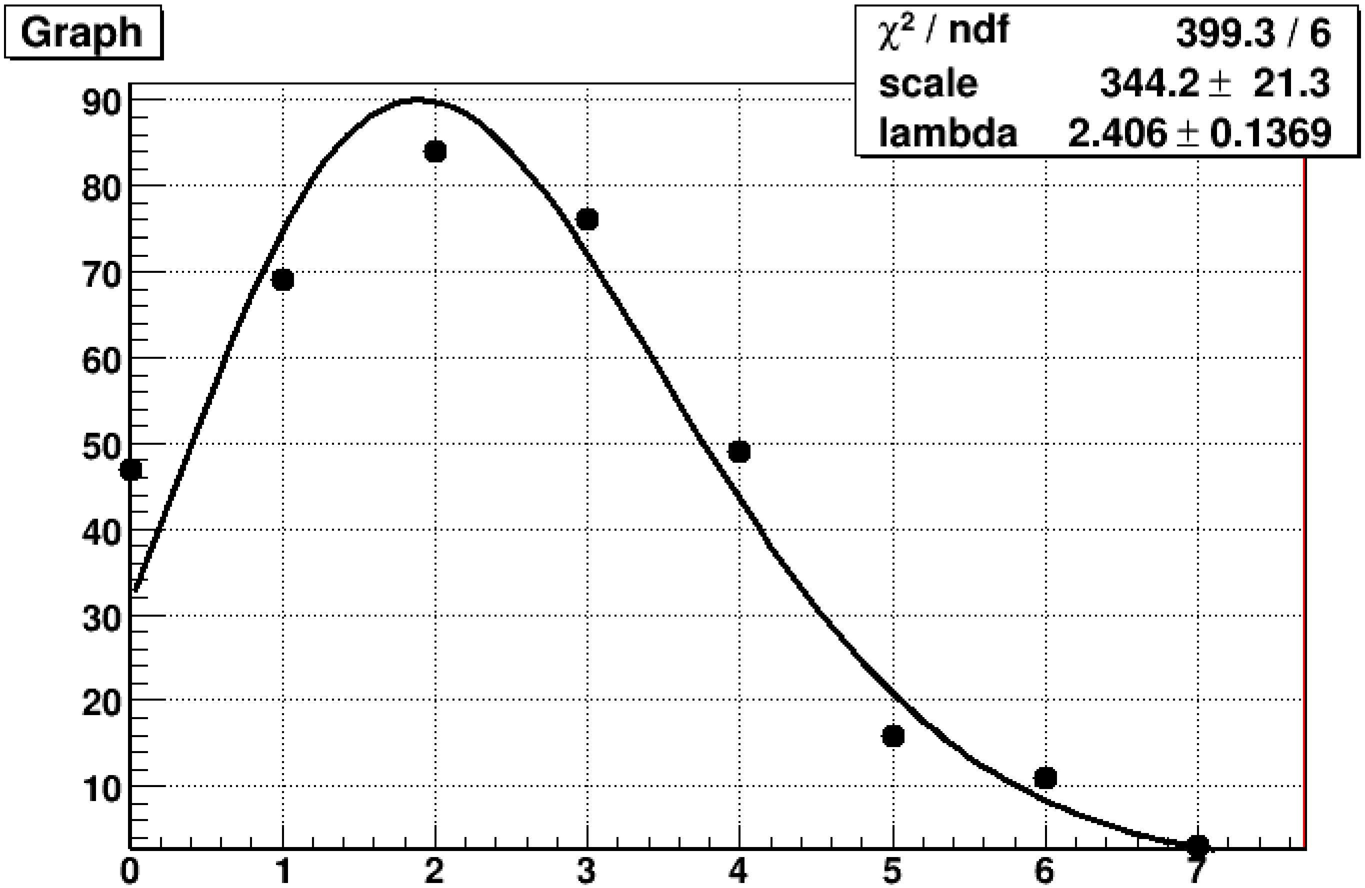
(r = 8, p = 1)

Dla poziomu istotności α = 1%

Otrzymujemy z tabeli $X^2_{0.99} = 16.81$

Nie ma podstaw do odrzucenia hipotezy.

Test X^2 i dopasowanie rozkładu Poissona



$$X^2 = \sum_k \frac{(n_k - n\,p_k)^2}{n\,p_k}$$

KONIEC WYKŁADU 9