



Komputerowa analiza danych doświadczalnych

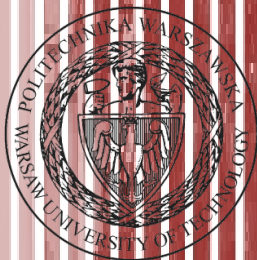
Wykład 10

26.04.2019

dr inż. Łukasz Graczykowski

lukasz.graczykowski@pw.edu.pl

Semestr letni 2018/2019



Przykłady funkcji wiarygodności i prawo kombinacji niepewności

Własności asymptotyczne



Prawo kombinacji niepewności

Prawo kombinacji niepewności

- Powróćmy do problemu z poprzedniego wykładu – wielokrotny pomiar tej samej wielkości z różną dokładnością, lub (równoważnie) problem pobierania próby z rozkładów normalnych o jednakowej wartości średniej λ i różnych, lecz znanych wariancjach σ_j

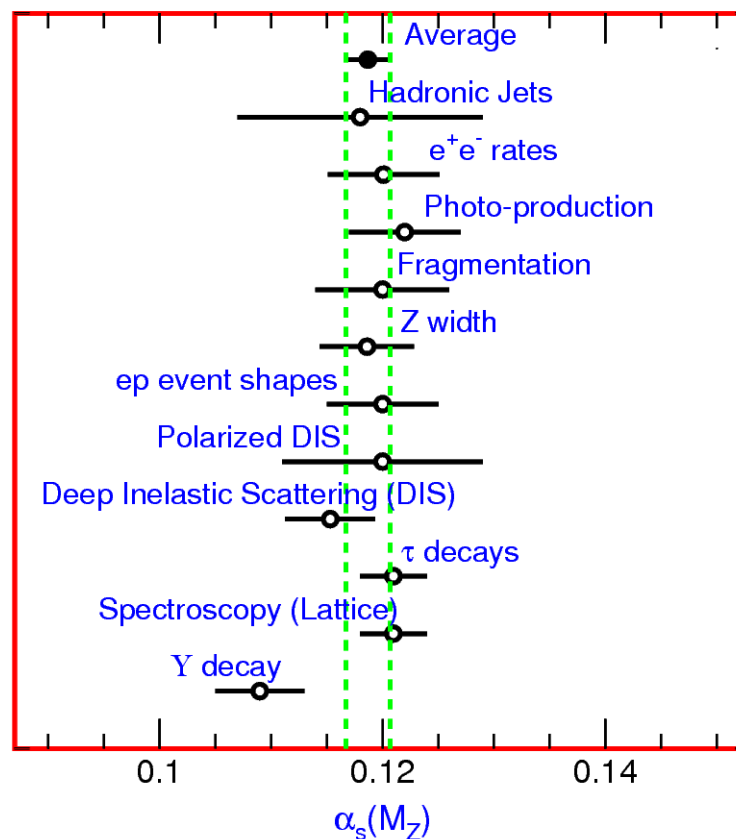


Figure 9.1: Summary of the value of $\alpha_s(M_Z)$ from various processes. The values shown indicate the process and the measured value of α_s extrapolated to $\mu = M_Z$. The error shown is the *total* error including theoretical uncertainties. The average quoted in this report which comes from these measurements is also shown. See text for discussion of errors.

Prawo kombinacji niepewności

- Jak pamiętamy, logarytmiczna funkcja wiarygodności i jej pochodna (równanie wiarygodności) dla takiego przypadku:

$$l = -\frac{1}{2} \sum_{j=1}^N \frac{(X^{(j)} - \lambda)^2}{\sigma_j^2} + \text{const}$$

$$l' = \frac{dl}{d\lambda} = \sum_{j=1}^N \frac{X^{(j)} - \lambda}{\sigma_j^2} = 0$$

- Możemy r. wiarygodności przekształcić:

$$l' = \sum_{j=1}^N \frac{X^{(j)}}{\sigma_j^2} - \sum_{j=1}^N \frac{\lambda}{\sigma_j^2} = \sum_{j=1}^N \frac{1}{\sigma_j^2} \left(\frac{\sum_{j=1}^N \frac{X^{(j)}}{\sigma_j^2}}{\sum_{j=1}^N \frac{1}{\sigma_j^2}} - \lambda \right) = 0$$

- I jego rozwiązanie:

$$S = \tilde{\lambda} = \sum_{j=1}^N \frac{X^{(j)}}{\sigma_j^2} / \sum_{j=1}^N \frac{1}{\sigma_j^2}$$

- Estymator S jest nieobciążonym estymatorem wartości średniej λ .
- Jaka jest niepewność (wariancja) tego estymatora?
 - W naszym przykładzie** – jaka jest niepewność wartości średniej stałej sprzężenia oddziaływania silnego na podstawie różnych pomiarów?

Przykład – estymacja rozkładu norm.

- oraz logarytmiczną funkcję wiarygodności:

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

- równania wiarygodności:

Wykład 9 - przypomnienie

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

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

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

- dla formalności powinniśmy jeszcze sprawdzić drugie pochodne...

Prawo kombinacji niepewności

- Jeśli teraz porównamy:

$$l' = \sum_{j=1}^N \frac{X^{(j)}}{\sigma_j^2} - \sum_{j=1}^N \frac{\lambda}{\sigma_j^2} = \sum_{j=1}^N \frac{1}{\sigma_j^2} \left(\frac{\sum_{j=1}^N \frac{X^{(j)}}{\sigma_j^2}}{\sum_{j=1}^N \frac{1}{\sigma_j^2}} - \lambda \right) = \boxed{\sum_{j=1}^N \frac{1}{\sigma_j^2}} (\tilde{\lambda} - \lambda)$$

- Ze wzorem z ostatniego wykładu:

$$l' = A(\lambda)(S - E(S)) = A(\lambda)(S - B(\lambda) - \lambda)$$

$$\text{dla } S = \tilde{\lambda} \text{ i } B(\lambda) = 0$$

$$l' = \boxed{A(\lambda)} (\tilde{\lambda} - \lambda)$$

- Widzimy, że estymator ten ma również **minimalną wariancję**. Jaka ona jest? Posłużmy się wzorem z ostatniego wykładu:

- Czyli:

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

$$\sigma^2(\tilde{\lambda}) = 1 / \sum_{j=1}^N \frac{1}{\sigma_j^2}$$

- Wzór ten określamy jako **prawo kombinacji niepewności**, lub (bardziej kolokwialnie) “**uśrednianie niepewności w kwadratach**”

Prawo kombinacji niepewności

- Możemy zauważyć, że jeżeli **wszystkie** pomiary mają jednakową dokładność, otrzymamy:

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

- Analogicznie jak na Wykładzie 7 i 8, gdy liczyliśmy wartość oczekiwaną i wariancję z próby za pomocą estymatorów
- Wniosek:** w zależności od potrzeb, możemy stosować podejście zarówno poprzez obliczanie estymatorów bezpośrednio (jeżeli potrafimy je wprowadzić), jak i poprzez metodę największej wiarygodności

Estymator wartości oczekiwanej

Próba losowa

- **wariancja średniej arytmetycznej:**

$$\sigma^2(\bar{X}) = E\{\bar{X} - E(\bar{X})\}^2 = E\left\{\left(\frac{X_1 + X_2 + \dots + X_n}{n} - \hat{x}\right)^2\right\}$$

$$= \frac{1}{n^2} E\{[(X_1 - \hat{x}) + (X_2 - \hat{x}) + \dots + (X_n - \hat{x})]^2\}$$

- **ponieważ elementy próby losowej są niezależne:**

$$\text{Cov}(X_i, X_j) = E\{(X_i - \hat{x})(X_j - \hat{x})\} = 0$$

- **to upraszczamy nawias:**

$$\begin{aligned} E\{[(X_1 - \hat{x}) + (X_2 - \hat{x}) + \dots + (X_n - \hat{x})]^2\} &= E\{(X_1 - \hat{x})^2 + (X_2 - \hat{x})^2 + \dots + (X_n - \hat{x})^2\} \\ &= E(X_1 - \hat{x})^2 + E(X_2 - \hat{x})^2 + \dots + E(X_n - \hat{x})^2 = \sigma^2(X_1) + \sigma^2(X_2) + \dots + \sigma^2(X_n) = n \sigma^2(X) \end{aligned}$$

- **zatem ostatecznie wariancja średniej arytmetycznej:**

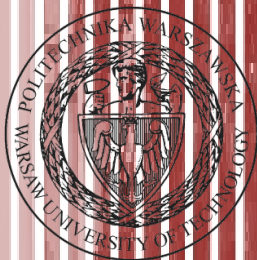
$$\sigma^2(\bar{X}) = \frac{1}{n^2} n \sigma^2(X) = \frac{1}{n} \sigma^2(X)$$

- **ponieważ wariancja rozkładu populacji jest jedną liczbą, to:**

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

Wykazaliśmy, że estymator jest **zgodny**

**PRZYPOMNIENIE
WYKŁAD 7 I 8**



Własności asymptotyczne estymatorów i funkcji wiarygodności

Własności asymptotyczne estymatorów

- Omówimy teraz niektóre własności funkcji wiarygodności i estymatorów największej wiarygodności dla dużych prób, czyli $N \rightarrow \infty$
- Estymator $S = \tilde{\lambda}$ jest zdefiniowany jako rozwiązanie równania wiarygodności:

$$l'(\lambda) = \sum_{j=1}^N \left(\frac{f'(X^{(j)}; \lambda)}{f(X^{(j)}; \lambda)} \right)_{\tilde{\lambda}} = 0 \qquad I(\lambda) = E(l'^2) = NE \left(\left(\frac{f'(x^{(j)}; \lambda)}{f(x^{(j)}; \lambda)} \right)^2 \right)$$

- Rozwijamy $l'(\lambda)$ na szereg Taylora wokół punktu $\lambda = \tilde{\lambda}$
- Czyli: $l'(\lambda) = l'(\tilde{\lambda}) + (\lambda - \tilde{\lambda})l''(\tilde{\lambda}) + \dots$
- Pierwszy wyraz z prawej strony znika (z uwagi na r. wiar.), drugą pochodną możemy zaś zapisać jako (w granicy $N \rightarrow \infty$):

$$l''(\tilde{\lambda}) = \sum_{j=1}^N \left(\frac{f'(X^{(j)}; \lambda)}{f(X^{(j)}; \lambda)} \right)'_{\tilde{\lambda}} \Rightarrow l''(\tilde{\lambda}) = NE \left(\left(\frac{f'(X^{(j)}; \lambda)}{f(X^{(j)}; \lambda)} \right)'_{\tilde{\lambda}} \right)$$

- Jeśli teraz posłużymy się wzorem (poprz wykład): $I(\lambda) = E(l'^2) = -E(l'')$

Własności asymptotyczne estymatorów

- W takim przypadku możemy zapisać:

$$l''(\tilde{\lambda}) = E(l''(\tilde{\lambda})) = -E(l'^2(\tilde{\lambda})) = -I(\tilde{\lambda}) = -1/b^2$$

analogicznie do

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

stała \rightarrow

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

- Innymi słowy:** zastąpiliśmy drugą pochodną logarytmicznej funkcji wiarygodności $l''(\tilde{\lambda})$, która jest funkcją próby, liczbą, która zależy jedynie od gęstości f i estymatora $\tilde{\lambda}$
- Czyli nasze rozwinięcie na szereg Taylora (pomijając wyższe wyrazy):
$$l'(\lambda) = (-1/b^2)(\lambda - \tilde{\lambda}) \Rightarrow l(\lambda) = (-1/2b^2)(\lambda - \tilde{\lambda})^2 + c$$
- Stałą c wyznaczamy podstawiając: $\lambda = \tilde{\lambda} \Rightarrow c = l(\tilde{\lambda})$
- I ostatecznie dostajemy ($k = \text{const}$):

$$l(\lambda) - l(\tilde{\lambda}) = -\frac{1}{2} \frac{(\lambda - \tilde{\lambda})^2}{b^2} \Rightarrow L(\lambda) = k \exp\left(-\frac{(\lambda - \tilde{\lambda})^2}{2b^2}\right)$$

- Funkcja wiarygodności (dla dużych N) ma więc rozkład normalny**

Własności asymptotyczne estymatorów

- Funkcja wiarygodności (dla dużych N) ma więc **rozkład normalny**

$$L(\lambda) = k \exp\left(-\frac{(\lambda - \tilde{\lambda})^2}{2b^2}\right)$$

$$\text{dla } \lambda = \tilde{\lambda}, \quad u(\tilde{\lambda}) = b \quad \lambda = \tilde{\lambda} \pm b$$

$$\text{otrzymujemy } -(l(\lambda) - l(\tilde{\lambda})) = \frac{1}{2}$$

- Ponieważ estymowaliśmy parametr λ mamy więc:

$$S = \tilde{\lambda}, \quad E(S) = \lambda$$

- Czyli $\tilde{\lambda}$ jest nieobciążonym estymatorem o minimalnej wariancji wynoszącej:

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

- Estymator ma wszystkie właściwości (nieobciążony, zgodny, z minimalną wariancją) **jedynie** w granicy $N \rightarrow \infty$
- Taki estymator nazywany jest estymatorem **asymptotycznie nieobciążonym**
- Analogicznie możemy powiedzieć, że funkcja wiarygodności jest **asymptotycznie normalna**

Własności asymptotyczne estymatorów

- Powiedzieliśmy sobie poprzednio, że funkcja wiarygodności jest miarą prawdopodobieństwa (rozkładem prawdopodobieństwa), ale dla parametrów badanego rozkładu prawd. danej cechy
- Zatem estymację parametru możemy przedstawić jako **wynik i niepewność typu A**:

$$\lambda = \tilde{\lambda}, \quad u(\lambda) = \sigma(\tilde{\lambda})$$

- Ponieważ funkcja wiarygodności asymptotycznie dąży do rozkładu normalnego, możemy powiedzieć, że wartość prawdziwa estymowanego parametru zawarta jest w przedziale:

$$\tilde{\lambda} - \sigma(\tilde{\lambda}) < \lambda_0 < \tilde{\lambda} + \sigma(\tilde{\lambda})$$

z prawdopodobieństwem 68,3% (“jeden sigma”)

- Podane związki możemy stosować jedynie **dla prób o dużym N** , dokładność przybliżeń zależy jednak również od rozpatrywanej funkcji gęstości – w każdym przypadku musimy zdecydować, czy powyższa procedura jest uprawniona do zastosowania

Przykład – wyznaczanie śr. czasu życia

- **Przykład:** chcemy zbadać średni czas τ życia jądra jakiegoś izotopu promieniotwórczego
- Prawdopodobieństwo że, istniejące w chwili $t=0$ jądro promieniotwórcze rozpadnie się w przedziale czasu od t do $t+dt$ wynosi:

$$f(t)dt = \frac{1}{\tau} \exp(-t/\tau) dt$$

- Badamy oczywiście pewną próbkę zawierającą wiele jąder – założmy, że zaobserwowaliśmy rozpady w chwilach: t_1, t_2, \dots, t_N
- Konstruujemy zatem odpowiednią funkcję wiarygodności:

$$L(\tau) = \frac{1}{\tau^N} \exp\left(-\frac{1}{\tau} \sum_{i=1}^N t_i\right) = \frac{1}{\tau^N} \exp\left(-\frac{N}{\tau} \bar{t}\right)$$

- Obliczamy logarytm i pierwszą pochodną:

$$l(\tau) = \ln L(\tau) = -\frac{N}{\tau} \bar{t} - N \ln \tau$$

$$l'(\tau) = \frac{N}{\tau} \left(\frac{\bar{t}}{\tau} - 1 \right) = \frac{N}{\tau^2} (\bar{t} - \tau)$$

Przykład – wyznaczanie śr. czasu życia

- Porównując to ze wzorem:

$$l' = A(\lambda)(S - E(S)) = A(\lambda)(S - B(\lambda) - \lambda)$$

$$\text{dla } S = \tilde{\lambda} \text{ i } B(\lambda) = 0$$

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

$$l'(\tau) = \frac{N}{\tau} \left(\frac{\bar{t}}{\tau} - 1 \right) = \frac{N}{\tau^2} (\bar{t} - \tau)$$

- Widzimy, że rozwiązaniem o największej wiarygodności jest:

$$\tilde{\tau} = \bar{t} \quad \sigma^2(\tilde{\tau}) = \tau^2 / N \Rightarrow \text{dla } \tau = \tilde{\tau} = \bar{t} \Rightarrow \sigma^2(\tilde{\tau}) = \bar{t} / N$$

- Wstawiając $\tau = \tilde{\tau} = \bar{t}$ do wzoru na l otrzymamy:

$$l(\tilde{\tau}) = l_{\max} = -N(1 + \ln \bar{t})$$

- Czyli:

$$-(l(\tau) - l(\tilde{\tau})) = N \left(\frac{\bar{t}}{\tau} + \ln \frac{\tau}{\bar{t}} - 1 \right)$$

- W tym wyrażeniu trudno jest rozpoznać asymptotyczną postać równania jak tutaj:

$$l(\lambda) - l(\tilde{\lambda}) = -\frac{1}{2} \frac{(\lambda - \tilde{\lambda})^2}{b^2} \Rightarrow L(\lambda) = k \exp \left(-\frac{(\lambda - \tilde{\lambda})^2}{2b^2} \right)$$

Przykład – wyznaczanie śr. czasu życia

- Czyli nie możemy od razu zastosować rozumowania, że wartość prawdziwa estymowanego mieści się w przedziale:

$$\tilde{\lambda} - \sigma(\tilde{\lambda}) < \lambda_0 < \tilde{\lambda} + \sigma(\tilde{\lambda})$$

z prawdopodobieństwem 68,3% (“1 sigma”)

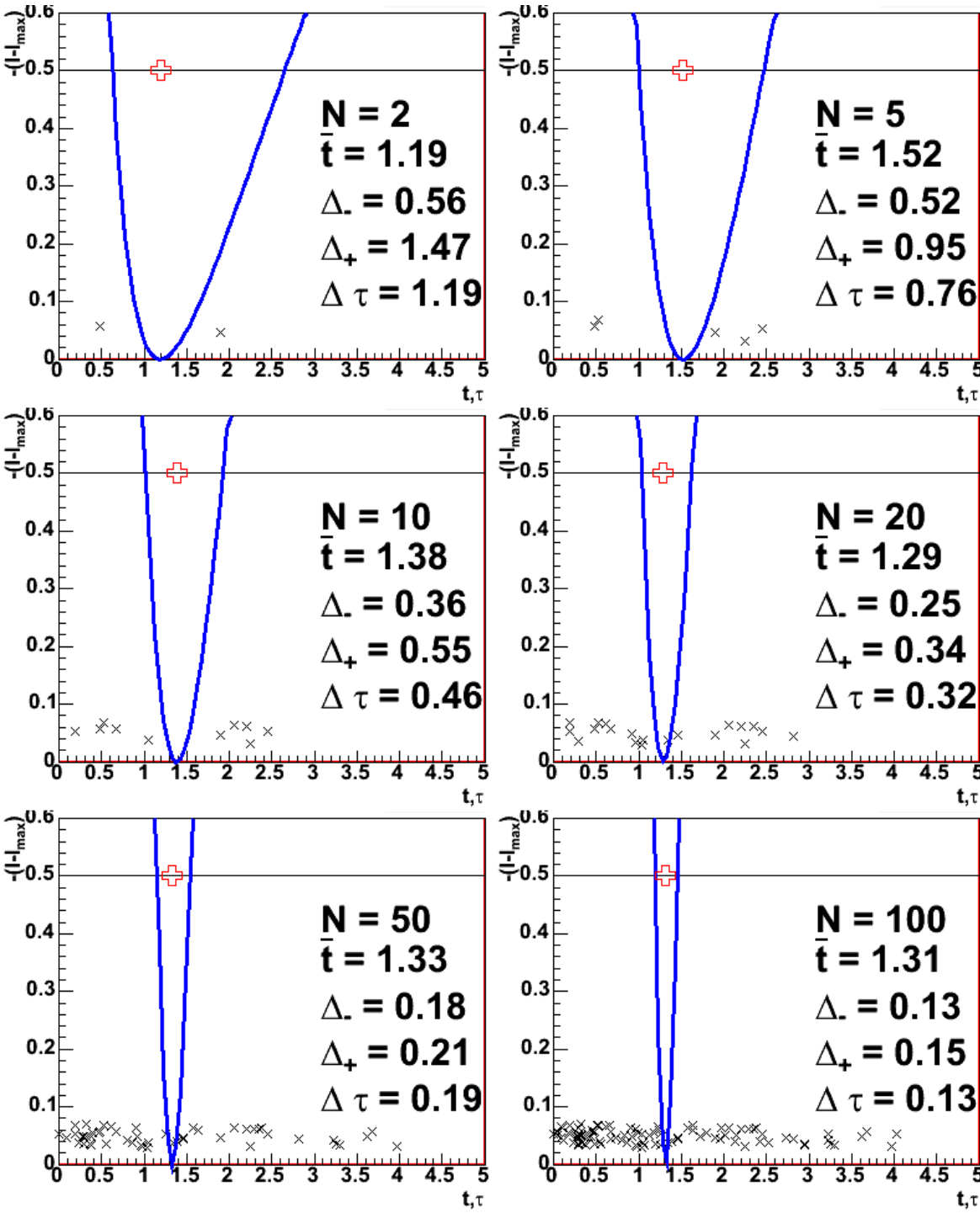
- Co zrobić? Jak żyć? :)
- Możemy jednak zdefiniować wartości τ_+ oraz τ_- takie, że:

$$\tau_+ = \tilde{\tau} + \Delta_+ \quad \tau_- = \tilde{\tau} + \Delta_-$$

- Które spełniają warunek: $-(l(\tau_{\pm}) - l(\tilde{\tau})) = \frac{1}{2}$
- Wzór ten definiuje **niepewności niesymetryczne (asymetryczne)**

- W granicy $N \rightarrow \infty$ oczekivalibyśmy: $\Delta_-, \Delta_+ \rightarrow u(\tilde{\tau}) = \sigma(\tilde{\tau})$
- Patrz rysunek na następnym slajdzie

Przykład – wyznaczanie śr. czasu życia

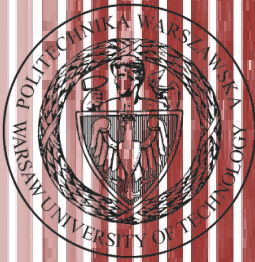


- Rysunek przedstawia funkcję:

$$-(l - l_{\max}) = -(l(\tau) - l(\tilde{\tau}))$$
- Punkty τ_{+} , τ_{-} znajdują się na przecięciu z prostą:

$$-(l(\tau) - l(\tilde{\tau})) = 1/2$$
- Krzyżem jest zaznaczony punkt:

$$\tilde{\tau} = \bar{t}$$
- Widzimy, że wraz ze zwiększaniem się N funkcja $-(l - l_{\max}) = -(l(\tau) - l(\tilde{\tau}))$
- Coraz bardziej przybiera kształt symetrycznej paraboli a niepewności asymetryczne coraz mniej się różnią od siebie
- Wniosek:** dla dużych prób f . wiaryg. ma rozkład Gaussa



Jednoczesna estymacja kilku parametrów

Jednoczesna estymacja kilku parametrów

- Jak już pamiętamy, w najogólniejszym przypadku mamy pewną ilość p parametrów $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_p)$, które chcemy estymować. Wtedy funkcja wiarygodności i równania wiarygodności:

$$L = \prod_{j=1}^N f(\mathbf{X}^{(j)}; \lambda) \qquad l = \ln L = \sum_{j=1}^N \ln f(\mathbf{X}^{(j)}; \lambda) \qquad \frac{\partial l}{\partial \lambda_i} = 0, \quad i = 1, 2, \dots, p$$

- Niestety, w przypadku wielu parametrów wszystko nam się komplikuje dodatkowo z uwagi na możliwe **korelacje pomiędzy parametrami**
- Analogicznie jak dla 1 parametru, rozwijamy funkcję wiarygodności:

$$L(X^{(1)}, X^{(2)}, \dots, X^{(j)}; \lambda) = \sum_{j=1}^N \ln f(X^{(j)}; \lambda)$$

na szereg Taylora wokół punktu:

- Korzystając ze znikania 1-szych pochodnych dostajemy wtedy:

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

- Gdzie macierz A :

$$\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p)$$
$$-A = \begin{pmatrix} \frac{\partial^2 l}{\partial \lambda_1^2} & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_p} \\ \frac{\partial^2 l}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 l}{\partial \lambda_2^2} & \cdots & \frac{\partial^2 l}{\partial \lambda_2 \partial \lambda_p} \\ \frac{\partial^2 l}{\partial \lambda_p \partial \lambda_1} & \frac{\partial^2 l}{\partial \lambda_p \partial \lambda_2} & \cdots & \frac{\partial^2 l}{\partial \lambda_p^2} \end{pmatrix}$$

Jednoczesna estymacja kilku parametrów

- W granicy $N \rightarrow \infty$ można zastąpić macierz A odpowiednimi wartościami oczekiwanymi (niezależnymi od próby) $B = E(A)$
 - de facto zamieniamy** $\lambda_1, \lambda_2, \dots, \lambda_p$ **na** $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$
- Jeśli zaniedbamy wyrazy wyższych rzędów w rozwinięciu na szereg Taylora, dostaniemy wówczas funkcję wiarygodności postaci:

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

- Analogicznie jak dla jednego parametru, mamy tutaj p -wymiarowy rozkład normalny z macierzą kowariancji (dla estymatorów)
- Wariancje estymatorów największej wiarygodności dane są przez elementy diagonalne macierzy C : $\sigma^2(\tilde{\lambda}_i) = c_{ii}$
- Elementy pozadiagonalne są kowariancjami poszczególnych par estymatorów: $\text{cov}(\tilde{\lambda}_j, \tilde{\lambda}_k) = c_{jk}$ $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_p$
- Możemy zdefiniować **współczynnik korelacji między estymatorami**: $\rho(\tilde{\lambda}_j, \tilde{\lambda}_k) = \frac{\text{cov}(\tilde{\lambda}_j, \tilde{\lambda}_k)}{\sigma(\tilde{\lambda}_j)\sigma(\tilde{\lambda}_k)}$

Jednoczesna estymacja kilku parametrów

- Analogicznie jak w przypadku z jednym parametrem, niepewnością estymacji (typu A) będzie pierwiastek kwadratowy z wariancji estymatora:

$$u(\tilde{\lambda}_i) = \sigma(\tilde{\lambda}_i) = \sqrt{c_{ii}}$$

- W przypadku 1D stwierdziliśmy, że za pomocą estymatora i jego niepewności możemy zdefiniować przedział obejmujący wartość prawdziwą estymowanego parametru z prawdopodobieństwem 68,3%. W przypadku wielowymiarowym przedział ten będzie określony również przez pełną macierz kowariancji (zależności między parametrami)
- Innymi słowy: obszar taki sprowadza się do elipsy kowariancji, analogicznej jak dla zmiennych losowych – czyli mamy **elipsy kowariancji dla estymowanych parametrów naszego rozkładu**
- Równanie **elipsoidy kowariancji** (p -wymiarowa przestrzeń):

$$g(\lambda) = 1 = 2 \{ l(\lambda) - l(\tilde{\lambda}) \} = (\lambda - \tilde{\lambda})^T B (\lambda - \tilde{\lambda})$$

- Dla $g(\lambda) = 1$ mamy obszar ufności z prawdopodobieństwem 68,3%



Jednoczesna estymacja kilku parametrów - przykład

Jednoczesna estymacja kilku parametrów

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

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

- Układ równań wiarygodności:

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

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

- Rozwiązanie:

$$\tilde{\lambda}_1 = \frac{1}{N} \sum_{j=1}^N X^{(j)}$$

$$\tilde{\lambda}_2 = \sqrt{\frac{\sum_{j=1}^N (X^{(j)} - \lambda_1)^2}{N}}$$

Jednoczesna estymacja kilku parametrów

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

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

$$B = \begin{pmatrix} N/\tilde{\lambda}_2^2 & 0 \\ 0 & 2N/\tilde{\lambda}_2^2 \end{pmatrix} \quad C = B^{-1} = \begin{pmatrix} \tilde{\lambda}_2^2/N & 0 \\ 0 & \tilde{\lambda}_2^2/2N \end{pmatrix}$$

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

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

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

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

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

Metoda najw. wiar. a programy do analizy

7.1 The Fit Method

The Fit method is implemented in ROOT for the histogram classes `TH1`, the sparse histogram classes, `THnSparse`, the graph classes, `TGraph`, `TGraph2D` and `TMultiGraph` for fitting a collection of Graphs with the same function.

7.1.1 The TH1::Fit Method

To fit a histogram programmatically, you can use the `TH1::Fit` method. Here is the signatures of `TH1::Fit` and an explanation of the parameters:

```
TFitResultPtr Fit(TF1 *function, Option_t *option, Option_t *goption,  
                  Axis_t xmin, Axis_t xmax)
```

- `function` a pointer to the fitted function (the fit model) object. One can also use the function name. This name may be one of ROOT pre-defined function names or a user-defined function. See the next paragraph for the list of pre-defined functions.
- `*option`: The second parameter is the fitting option. Here is the list of fitting options:
 - “`W`” Set all weights to 1 for non empty bins; ignore error bars
 - “`WW`” Set all weights to 1 including empty bins; ignore error bars
 - “`I`” Use integral of function in bin instead of value at bin center
 - “`L`” Use log likelihood method (default is chi-square method). To be used when the histogram represents counts
 - “`WL`” Weighted log likelihood method. To be used when the histogram has been filled with weights different than 1.
 - “`P`” Use Pearson chi-square method, using expected errors instead of the observed one given by `TH1::GetBinError` (default case). The expected error is instead estimated from the the square-root of the bin function value.

Metoda najw. wiar. a programy do analizy

Example: Likelihood fit

- *Use the pdf of observables and maximize the product of likelihoods.*
 - In practice: minimize the $-2 \times \text{logarithm}$ of the likelihoods

• Binned data

- Use Poisson pdf for bin content
- **Example:** data normally distributed

$$\mathcal{L} = \prod_{i=\text{bins}} \text{Poisson} \left[N_i^{\text{obs}}, A \exp \left(-\frac{1(x_i - \bar{x})^2}{2\sigma^2} \right) \right]$$

• Unbinned data

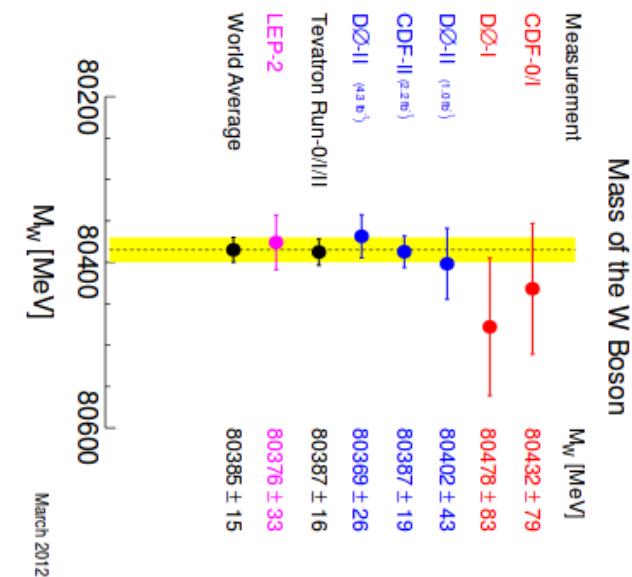
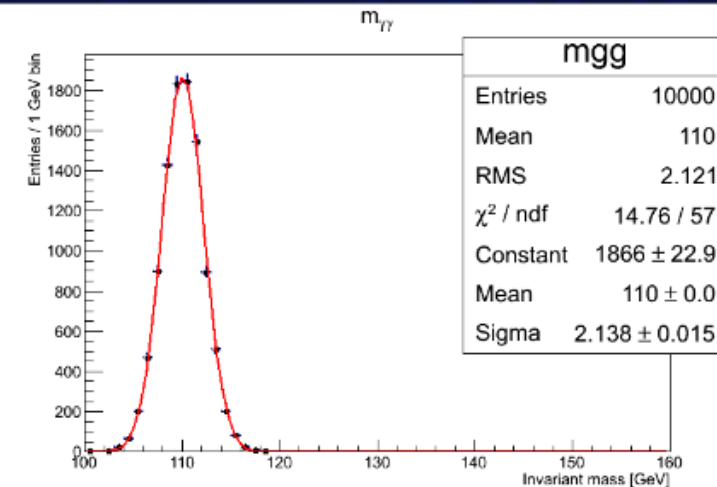
- **Example:** normally distributed points about α

$$\mathcal{L} = \prod_{i=\text{measurements}} \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left(-\frac{1(x_i - \alpha)^2}{2\sigma_i^2} \right)$$

$$-2 \ln \mathcal{L} = \sum_{i=\text{measurements}} \left[\frac{(x_i - \alpha)^2}{\sigma_i^2} + 2 \ln(\sqrt{2\pi}\sigma_i) \right]$$

- It looks familiar, doesn't it?

- **1σ uncertainty:** increase of $-2\ln\mathcal{L}$ by 1



Metoda najw. wiar. a programy do analizy

Simple fits

```
myh2->Fit("gaus","", "", -2., 2.);
```



```
myh2->GetFunction("gaus")->Draw();
```

Retrieve a pointer to the fitted TF1

The pointer can be used to manipulate the function: plotting, changing style, inspecting parameters...

```
TFitResultPtr fitres = myh2->Fit("gaus","S");
```

```
fitres->Parameter(2);
```

```
fitres->ParError(2);
```

```
fitres->Chi2();
```

```
fitres->Ndf();
```

- Some predefined fit functions available:

- Exponential
- Gaussian
- Landau
- Polynomial (up to 9th degree)

- Most common options:

- "L" likelihood fit
- "WL" likelihood fit with weighed points
- "Q" quiet mode: do not print info on screen
- "S" save the results in a TFitResultPtr object.





KONIEC