Analiza danych

Mariusz Przybycień

Wydział Fizyki i Informatyki Stosowanej
Akademia Górniczo-Hutnicza

Wykład 4

Wnioskowanie statystyczne

- Wnioskowanie statystyczne to proces matematyczny, którego celem jest wydobycie istotnych informacji z pomiarów, umożliwiających uzyskanie charakterystyk badanych układów, ustalenia najlepszych parametrów modeli lub określenia wiarygodności hipotez i modeli.
- Ogólnie rozróżniamy dwa podejścia do wnioskowania statystycznego oparte na:
  - częstościowej interpretacji p-twa:
    - estymacja punktowa to oszacowanie wartości wybranych obserwabli, parametrów rozkładów i/lub modeli wraz z niepewnościami,
    - estymacja przedziałowa to oszacowanie przedziału ufności w którym dana obserwabla (parametr rozkładu lub modelu), znajdzie się z określonym p-twem jeśli eksperyment byłby powtarzany nieskończenie wiele razy,
    - weryfikacja hipotez w oparciu o testy statystyczne skonstruowane w oparciu o probabilistyczne modele danych,
    - klastrowanie (uczenie bez nadzoru) i klasyfikacja (uczenie z nadzorem) danych - podział danych na grupy/klastry na podstawie wybranych cech z zastosowaniem metod uczenia maszynowego.
  - podejściu bayesowskim (dwie kategorie: wybór modelu lub szacowanie parametrów).

Wnioskowanie statystyczne - podstawowe pojęcia

- Populacja to zbiór wszystkich elemetów posiadających daną cechę (lub cechy) - zbiór osób zamieszkujących na terenie Polski, zbiór wszystkich gwiazd w Drodze Mlecznej, wybrana klasa przypadków w zderzeniach proton-proton, ...
- Podpopulacja to zbiór elementów danej populacji posiadających dodatkową wspólną cechę - w farmacji: na dany lek mogą różnie reagować różne grupy etniczne lub posiadające specyficzne mutacje genów; w HEP: możemy być zainteresowani własnościami (rozkład pędu, pseudorapidity, itp.) różnych typów cząstek naładowanych (piony, kaony, ...) produkowanych w określonych typach zderzeń (rodzaj zderzanych cząstek, energia w układzie środka masy, itp.
- W praktyce dysponujemy jedynie próbką losową - powinna być reprezentatywna dla całej populacji (nieobciążona), tzn. odzwierciedlać wszystkie cechy i związki w niej wytępujące.
- Próbka obciążona to próbka w której brakuje pewnych klas przypadków - teleskop może nie być w stanie dostrzec gwiazd, które nie są wystarczająco jasne, albo których widmo jest zbyt bardzo przesunięte ku czerwieni; w HEP: ograniczona akceptancja detektora (kinematyczna, geometryczna), tryger, wydajność detekcji i trygera.
- Próbka jest prosta jeśli wszystkie występujące w niej zmienne są niezależne - w HEP ten warunek najczęściej jest spełniowy, choć nie zawsze (np. STAR TPC).

Estymacja parametrów

- W ogólności rozważamy wektor losowy $\vec{\mathrm{X}}$ (n-tupel) oraz n-elementową próbkę losową prostą złożoną z n realizacji $\vec{\mathrm{X}}$ (przypadków w analizie HEP).

  Na tej podstawie chcemy oszacować nieznane wartości m parametrów $\vec{\theta}$ modelu M, który opisuje naszą wiedzę na temat badanego zjawiska, procesu, itp.
- Dla uproszczenia załóżmy, że wektor $\vec{\mathrm{X}} = (\mathrm{x})$ jest jednoelementowy. Dysponujemy więc n-elementową próbką prostą $\vec{x}_n = (x_1, x_2, ..., x_n)$ reprezentującą wyniki n pomiarów wielkości x, które zgodnie z naszym modelem zależą od m parametrów $\vec{\theta} = (\theta_1, \theta_2, ..., \theta_m)$.
- Dowolną funkcję $f(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ nazywamy statystyką.
- Estymatorem parametru θ nazywamy każdą statystykę $\hat{\theta}_n(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$, której wartości jesteśmy skłonni przyjąć jako oszacowanie tego parametru.
- Estymata parametru θ_m to wartość estymatora $\hat{\theta}_m$ jaką przyjmuje dla konkretnej próbki losowej.

Własności estymatorów

- Estymator $\hat{\theta}_n$ parametru θ nazywamy zgodnym, jeśli zachodzi:

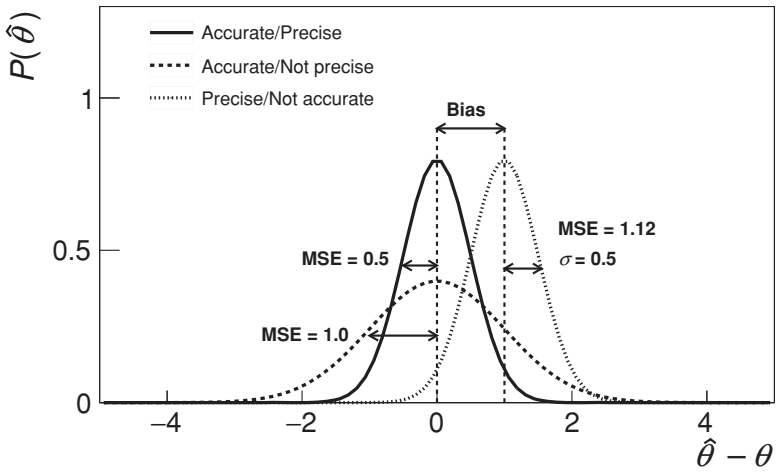
$$\lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\theta}_n - \theta\right| \leqslant \varepsilon\right) = 1 \quad \Leftrightarrow \quad \lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\theta}_n - \theta\right| > \varepsilon\right) = 0$$

- Estymator, którego wariancja dąży do zera dla liczebności próby dążącej do ∞, jest estymatorem zgodnym (na podstawie nierówności czebyszewa):

$$\mathrm{P}\left(\left|\hat{\theta}_n - \theta\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2_{\hat{\theta}_n}}{\varepsilon^2} \quad \Rightarrow \quad \left(\sigma^2_{\hat{\theta}_n} \xrightarrow[n\to\infty]{} 0 \quad \Rightarrow \quad \lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\theta}_n - \theta\right| \geqslant \varepsilon\right) = 0\right)$$

- Estymator nazywamy nieobciążonym jeśli jego warość oczekiwana jest równa wielkości estymowanej, $\mathcal{E}[\hat{\theta}_n] = \theta$.
- Jeśli $\mathcal{E}[\hat{\theta}_n] < \infty$, ale $\mathcal{E}[\hat{\theta}_n] \neq \theta$, to estymator nazywamy obciążonym, natomiast różnicę $b_n(\hat{\theta}) = \mathcal{E}[\hat{\theta}_n] - \theta$ nazywamy obciążeniem estymatora.
- Jeśli $b_n(\hat{\theta}) \underset{n\to\infty}{\longrightarrow} 0$, to estymator nazywamy asymptotycznie nieobciążonym.
- Średni błąd kwadratowy (MSE) estymatora:

$$\begin{aligned} \mathrm{MSE}(\hat{\theta}) &= \mathcal{E}[(\hat{\theta} - \theta)^2] = \\ &= \mathcal{V}[\hat{\theta} - \theta] + (\mathcal{E}[\hat{\theta} - \theta])^2 = \\ &= \mathcal{V}[\hat{\theta}] + b_n^2(\hat{\theta}) \end{aligned}$$



Podstawowe estymatory punktowe

- Nieobciążonym estymatorem wartości oczekiwanej jest średnia arytmetyczna:

$$\mathcal{E}[\overline{x}] = \mathcal{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathcal{E}[\mathrm{x}_i] = \frac{1}{n}\, n\, \mu = \mu$$

- Wariancja średniej arytmetycznej:

$$\mathcal{V}[\overline{x}] = \mathcal{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2\right] - \mathcal{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right]^2 = \frac{1}{n^2}(nm_2 + n(n-1)\mu^2) - \mu^2 = \frac{\sigma^2}{n}$$

- Nieobciążony estymator wariancji: $\mathrm{s}_\mathrm{x}^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$
- Wariancja estymatora wariancji:

$$\mathcal{V}[\mathrm{s}_\mathrm{x}^2] = \left\langle \mathrm{s}_\mathrm{x}^4 \right\rangle - \left\langle \mathrm{s}_\mathrm{x}^2 \right\rangle^2 = \frac{1}{n}\left\langle (\mathrm{x}-\mu)^4 \right\rangle - \frac{n-3}{n(n-1)}\mathcal{V}[\mathrm{x}]$$

- Estymator wariancji estymatora wariancji:

$$\mathrm{s}^2_{\mathrm{s}^2_x} = \frac{n}{(n-2)(n-3)}\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^4 - \frac{n^2-3}{n^2}\mathrm{s}_\mathrm{x}^4\right) \approx \frac{\sum_{i=1}^{n}\left((x_i-\overline{x})^2 - \mathrm{s}_\mathrm{x}^2\right)^2}{(n-2)(n-3)}$$

- Nieobciążony estymator kowariancji: $\mathrm{R} = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})$

Funkcja wiarogodności

- Niech będzie dana próba losowa prosta o liczebności n z rozkładu $f(x;\theta)$. Funkcją wiarogodności dla próby $x_i,\ (i=1,...,n)$, nazywamy wielkość:

$$\mathcal{L}(\vec{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

  Funkcja wiarogodności to nie to samo co łączna gęstość p-twa $L(\vec{x};\theta)$.

- Twierdzenie Cramera-Rao: Minimalna wartość wariancji $\mathcal{V}_{\min}$ dowolnego nieobciążonego estymatora parametru θ dana jest przez:

$$\mathcal{V}[\hat{\theta}] \geqslant \mathcal{V}_{\min}[\hat{\theta}] = \left\{ \mathcal{E}\left[\left(\frac{\partial}{\partial\theta} \ln L(\vec{x};\theta) \right)^2 \right] \right\}^{-1} = \left\{ \mathcal{E}\left[-\frac{\partial^2}{\partial\theta^2} \ln L(\vec{x};\theta) \right] \right\}^{-1}$$

  Uzasadnienie ostatniej równości w powyższym wzorze:

$$\int_{-\infty}^{\infty} L(\vec{x};\theta)d\vec{x} = 1 \quad \overset{\partial/\partial\theta}{\Longrightarrow} \quad 0 = \int_{-\infty}^{\infty} \frac{\partial \ln L(\vec{x};\theta)}{\partial\theta} L(\vec{x};\theta)d\vec{x} = \mathcal{E}\left[\frac{\partial \ln L(\vec{x};\theta)}{\partial\theta} \right]$$

$$\overset{\partial/\partial\theta}{\Longrightarrow} \quad \int_{-\infty}^{\infty} \left(\frac{\partial \ln L(\vec{x};\theta)}{\partial\theta} \right)^2 L(\vec{x};\theta)d\vec{x} = -\int_{-\infty}^{\infty} \frac{\partial^2 \ln L(\vec{x};\theta)}{\partial\theta^2} L(\vec{x};\theta)d\vec{x}$$

Twierdzenie Cramera-Rao

- W przypadku m parametrów $\vec{\theta} = (\theta_1, ..., \theta_m)$, nierówność Cramera-Rao ma charakter macierzowy:

$$\mathrm{V}_{\min}\left[\hat{\vec{\theta}}\right] = \left\{\mathcal{E}\left[\frac{\partial \ln L(\vec{x};\vec{\theta})}{\partial\theta_i}\cdot\frac{\partial \ln L(\vec{x};\vec{\theta})}{\partial\theta_j}\right]\right\}^{-1} = \left\{\mathcal{E}\left[-\frac{\partial^2 \ln L(\vec{x};\theta)}{\partial\theta_i\partial\theta_j}\right]\right\}^{-1}$$

Twierdzenie Cramera-Rao stwierdza, że $\mathrm{V} - \mathrm{V}_{\min}$ gdzie $\mathrm{V}_{ij} = \mathrm{cov}\left(\hat{\theta}_i, \hat{\theta}_j\right)$, jest macierzą dodatnie półokreśloną, w szczególności

$$\mathcal{V}\left[\hat{\theta}_i\right] \geqslant [\mathrm{V}_{\min}]_{ii}$$

$\mathrm{V}_{\min}$ wykorzystujemy jako przybliżenie macierzy kowariancji wstawiając w drugich pochodnych zamiast parametrów ich estymaty.

- W przypadku dyskretnego rozkładu p-twa mamy ($\vec{k} = (k_1, k_2, ..., k_n)$):

$$\mathcal{L}(\vec{k};\theta) = \prod_{i=1}^{n} P_{k_i}(\theta) \quad \Rightarrow \quad \mathcal{V}[\hat{\theta}] \geqslant \mathcal{V}_{\min}[\hat{\theta}] = \left\{\sum_{k_1,k_2,\ldots,k_n} L(\vec{k};\theta)\left(\frac{\partial}{\partial\theta}\ln L(\vec{k};\theta)\right)^2\right\}^{-1}$$

$$\mathcal{V}[\hat{\theta}] \geqslant \mathcal{V}_{\min}[\hat{\theta}] = \left\{n\mathcal{E}\left[\left(\frac{\partial}{\partial\theta}\ln P_k(\theta)\right)^2\right]\right\}^{-1} = -\left\{n\mathcal{E}\left[\frac{\partial^2}{\partial\theta^2}\ln P_k(\theta)\right]\right\}^{-1}$$

Minimalna wariancja estymatora wariancji rozkładu Gaussa

- Wyznaczenie $\mathcal{V}_{\min}\left[\hat{\sigma}^2\right]$ dla rozkładu Gaussa:

$$\begin{aligned}
\ln \mathcal{N}(x;\mu,\sigma) &= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \frac{(x-\mu)^2}{2\sigma^2} \\
\frac{\partial}{\partial(\sigma^2)}\ln \mathcal{N}(x;\mu,\sigma) &= -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2} \\
\frac{\partial^2}{\partial(\sigma^2)^2}\ln \mathcal{N}(x;\mu,\sigma) &= \frac{1}{2(\sigma^2)^2} - \frac{(x-\mu)^2}{(\sigma^2)^3}
\end{aligned}$$

  Z tw. Cramera-Rao otrzymujemy:

$$\mathcal{V}_{\min}\left[\hat{\sigma}^2\right] = \left(-n\int_{-\infty}^{\infty}\mathcal{N}(x;\mu,\sigma)\frac{\partial^2}{\partial(\sigma^2)^2}\ln\mathcal{N}(x;\mu,\sigma)dx\right)^{-1} = \frac{2}{n}\sigma^4$$

- Podobnie dla dyspersji wyznaczamy:

$$\mathcal{V}_{\min}\left[\hat{\sigma}\right] = \left(-n\int_{-\infty}^{\infty}\left(\frac{1}{\sigma^2} - 3\frac{(x-\mu)^2}{\sigma^4}\right)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right)^{-1} = \frac{\sigma^2}{2n}$$

Minimalna wariancja estymatora parametru rozkładu Poissona

- Parametr μ rozkładu Poissona to jednocześnie jego wartość oczekiwana i wariancja. Który z estymatorów - $\overline{\mathrm{x}}$ czy s^2_{x} - jest bardziej efektywny?

$$\ln \mathcal{P}_k(\mu) = k \ln \mu - \mu - \ln k!$$
$$\frac{\partial}{\partial \mu} \ln \mathcal{P}_k(\mu) = \frac{k}{\mu} - 1$$

Z tw. Cramera-Rao otrzymujemy:

$$\mathcal{V}_{\min}\left[\hat{\mu}\right] = \left(n \sum_{k=0}^{\infty} \mathcal{P}_k(\mu) \left(\frac{\partial}{\partial \mu} \ln \mathcal{P}_k(\mu) \right)^2 \right)^{-1} = \left(\frac{n}{\mu^2} \sum_{k=0}^{\infty} (k-\mu)^2 \mathcal{P}_k(\mu) \right)^{-1} = \frac{\mu}{n}$$

- Ponieważ

$$\mathcal{V}\left[s^2_{\mathrm{k}}\right] = \frac{1}{n}\langle(\mathrm{k}-\langle\mathrm{k}\rangle)^4\rangle - \frac{n-3}{n(n-1)}\mathcal{V}^2[\mathrm{k}] = \frac{1}{n}(\mu+3\mu^2) - \frac{n-3}{n(n-1)}\mu^2 = \frac{\mu}{n} + \frac{2\mu^2}{n-1}$$

więc efektywność tego estymatora to $\dfrac{\frac{\mu}{n}}{\frac{\mu}{n} + \frac{2\mu^2}{n-1}} = \dfrac{1}{1 + \frac{2n}{n-1}\mu} \xrightarrow[n\to\infty]{} \dfrac{1}{1+2\mu}$

Czyli dla $\mu \neq 0$ efektywność tego estymatora jest zawsze mniejsza od jedności.

- Najefektywniejszym estymatorem parametru μ jest średnia $\overline{\mathrm{x}}$.

Metoda Największej Wiarogodności (MNW)

- Zasada największej wiarogodności: Za estymatę nieznanego parametru θ należy przyjąć taką wartość $\hat{\theta}$, dla której funkcja wiarogodności osiąga maksimum:

$$\mathcal{L}(\vec{x};\hat{\theta}) = \max$$

a więc: $\dfrac{\partial}{\partial\theta}\mathcal{L}(\vec{x};\hat{\theta}) = 0$ przy warunku $\left.\dfrac{\partial^2}{\partial\theta^2}\mathcal{L}(\vec{x};\hat{\theta})\right|_{\theta=\hat{\theta}} < 0$

- Często jest wygodniej posługiwać się logarytmem $\mathcal{L}$:

$$\frac{\partial}{\partial\theta}\ln\mathcal{L}(\vec{x};\hat{\theta}) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(x_i;\hat{\theta}) = 0 \text{ przy warunku } \left.\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\ln f(x_i;\hat{\theta})\right|_{\theta=\hat{\theta}} < 0$$

- W przypadku m parametrów $\vec{\theta} = (\theta_1, ..., \theta_m)$ musimy rozwiązać układ m rownań:

$$\frac{\partial}{\partial\theta_j}\ln\mathcal{L}(\vec{x};\hat{\theta}) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta_j}\ln f(x_i;\hat{\theta}) = 0$$

przy warunku ujemnej określoności macierzy: $\left.\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta_j\partial\theta_k}\ln f(x_i;\vec{\theta})\right|_{\vec{\theta}=\hat{\vec{\theta}}}$

Własności estymatorów MNW

- Przykład: Wyznaczenie estymatora parametrów τ i λ rozkładu wykładniczego.

$$\mathcal{L}(t;\tau) = \prod_{i=1}^{n} \frac{1}{\tau} \exp\left(-\frac{t_i}{\tau}\right) \quad \Rightarrow \quad \ln \mathcal{L} = -n \ln \tau - \frac{1}{\tau} \sum_{i=1}^{n} t_i$$

$$\frac{\partial \ln \mathcal{L}}{\partial \tau} = -\frac{n}{\tau} + \frac{1}{\tau^2} \sum_{i=1}^{n} t_i \quad \Rightarrow \quad -\frac{n}{\hat{\tau}} + \frac{1}{\hat{\tau}} \sum_{i=1}^{n} t_i \quad \Rightarrow \quad \hat{\tau} = \frac{1}{n} \sum_{i=1}^{n} t_i = \bar{\mathrm{t}}$$

Podobnie wyznaczamy estymator parametru λ:

$$\mathcal{L}(t;\lambda) = \prod_{i=1}^{n} \lambda \exp\left(-\lambda t_i\right) \quad \Rightarrow \quad \ln \mathcal{L} = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i$$

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^{n} t_i \quad \Rightarrow \quad \frac{n}{\hat{\lambda}} + \sum_{i=1}^{n} t_i \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{\bar{\mathrm{t}}}$$

Widać, że estymator jest niezmienniczy względem transformacji $\alpha(\xi) = 1/\xi$.

- Jest to ogólna cecha estymatorów MNW, tzn. jeśli znamy estymator parametru θ, a interesuje nas funkcja tego parametru $\alpha(\theta)$, to zakładając że $\partial\alpha/\partial\theta \neq 0$, mamy:

$$0 = \frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial \alpha} \frac{\partial \alpha}{\partial \theta} \quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \alpha} = 0$$

W ogólności $\mathcal{E}[\alpha(\hat{\theta})] \neq \alpha(\mathcal{E}[\hat{\theta}])$ (z wyjątkiem przekształcenia liniowego), dlatego oczekujemy, że estymatory MNW są z reguły obciążone.

Szacowanie błędów estymatorów MNW

- Wariancja estymatora MNW: $\mathcal{V}[\hat{\theta}] = \int \left(\hat{\theta}(\vec{x}) - \langle\hat{\theta}\rangle\right)^2 L(\vec{x};\theta)\, d\vec{x}$

  Wariancja przyjmuje więc postać funkcji ocenianego parametru. Aby uzyskać wartść liczbową, musimy wstawić jego estymatę (jest to więc przybliżenie).

- W przypadku m parametrów musimy skonstruować macierz kowariancji:

$$\mathrm{V}\left[\hat{\theta}_i, \hat{\theta}_j\right] = \int \left(\hat{\theta}_i(\vec{x}) - \langle\hat{\theta}_i\rangle\right)\left(\hat{\theta}_j(\vec{x}) - \langle\hat{\theta}_j\rangle\right) L(\vec{x};\vec{\theta})\, d\vec{x}$$

- Przykład: Wariancja estymatora parametru τ rozkładu wykładniczego:

$$\mathcal{V}[\hat{\tau}] = \int_0^\infty \left(\frac{1}{n}\sum_{i=1}^n t_i - \tau\right)^2 \prod_{j=1}^n \frac{1}{\tau}\exp\left(-\frac{t_j}{\tau}\right) d\vec{t} =$$

$$= \frac{1}{\tau^n}\left(\int_0^\infty \left(\frac{1}{n}\sum_{i=1}^n t_i\right)^2 \prod_{j=1}^n \exp\left(-\frac{t_j}{\tau}\right) d\vec{t} - \right.$$

$$\left. -2\tau\int_0^\infty \left(\frac{1}{n}\sum_{i=1}^n t_i\right) \prod_{j=1}^n \exp\left(-\frac{t_j}{\tau}\right) d\vec{t} + \tau^2 \int_0^\infty \prod_{j=1}^n \exp\left(-\frac{t_j}{\tau}\right) d\vec{t}\right) =$$

$$= \frac{1}{\tau^n}\left(\frac{1}{n^2} n 2\tau^3 \tau^{m-1} + \frac{1}{n^2} n(n-1)\tau^2\tau^2\tau^{n-2} - \frac{2\tau}{n} n\tau^2\tau^{n-1} + \tau^2\tau^n\right) = \frac{\tau^2}{n}$$

Szacowanie błędów estymatorów MNW

- W przypadku dużej próbki danych możemy znaleźć estymator wariancji estymatora parametru MNW korzystając z nierówności Cramera-Rao:

$$\hat{\mathrm{V}}\left[\hat{\theta}_i, \hat{\theta}_j\right] = -\left(\left.\frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j}\right|_{\vec{\theta}=\hat{\vec{\theta}}}\right)^{-1}$$

- Przykład: Estymator wariancji estymatora parametru rozkładu wykładniczego:

$$\hat{\mathcal{V}}[\hat{\tau}] = \left(\left.\frac{\partial^2 \ln \mathcal{L}}{\partial\tau^2}\right|_{\tau=\hat{\tau}}\right) = \frac{\hat{\tau}^2}{n} \qquad \hat{\mathcal{V}}[\hat{\lambda}] = \left(\left.\frac{\partial^2 \ln \mathcal{L}}{\partial\lambda^2}\right|_{\lambda=\hat{\lambda}}\right) = \frac{\hat{\lambda}^2}{n}$$

- Obserwacja: Estymatory MNW są zgodne.
- Zachowanie funkcji wiarogodności jako funkcji parametru θ w okolicach estymaty MNW tego parametru:

$$\ln \mathcal{L}(\theta) = \underbrace{\ln \mathcal{L}(\hat{\theta})}_{\ln \mathcal{L}_{\max}} + \underbrace{\left.\frac{\partial \ln \mathcal{L}}{\partial\theta}\right|_{\theta=\hat{\theta}}}_{=0} + \frac{1}{2}\left.\frac{\partial \ln \mathcal{L}}{\partial\theta^2}\right|_{\theta=\hat{\theta}} (\theta-\hat{\theta})^2 + \mathcal{O}\left((\theta-\hat{\theta})^3\right)$$

$$\ln \mathcal{L}(\theta) \approx \ln \mathcal{L}_{\max} - \frac{(\theta-\hat{\theta})^2}{2\hat{\mathcal{V}}[\hat{\theta}]} \quad \Rightarrow \quad \mathcal{L}(\theta) \approx \mathcal{L}_{\max} \exp\left[-\frac{(\theta-\hat{\theta})^2}{2\hat{\mathcal{V}}[\hat{\theta}]}\right]$$

$\mathcal{L}(\theta)$ ma w przybliżeniu w okolicach maksimum rozkład Gaussa.

Szacowanie estymatorów MNW i ich błędów

- Przykład: Rozważmy pomiar średniego czasu życia $\tau = 1/\lambda$ (cząstki lub jądra) w eksperymencie w którym maksymalny czas pomiaru wynosi T.
  - Musimy zastosować obcięty rozkład wykładniczy, w którym $0 < t < T$:

$$f(t;\lambda) = \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda T}} \quad \Rightarrow \quad \langle t \rangle = \frac{\lambda}{1 - e^{-\lambda T}} \int_0^T t e^{-\lambda t} dt = \frac{1 - (1 + \lambda T) e^{-\lambda T}}{\lambda (1 - e^{-\lambda T})}$$

  Stosując metodę momentów możemy napisać: $\bar{t} = \langle t \rangle$ i wyznaczyć stąd $\hat{\tau} = 1/\hat{\lambda}$.

  - Zakładając, że każdy z pomiarów charakteryzuje się pewnym minimalnym, $t_i^{\min}$, i pewnym maksymalnym, $t_i^{\max}$, dostępnym czasem pomiaru, otrzymujemy dla pojedynczego rozpadu:

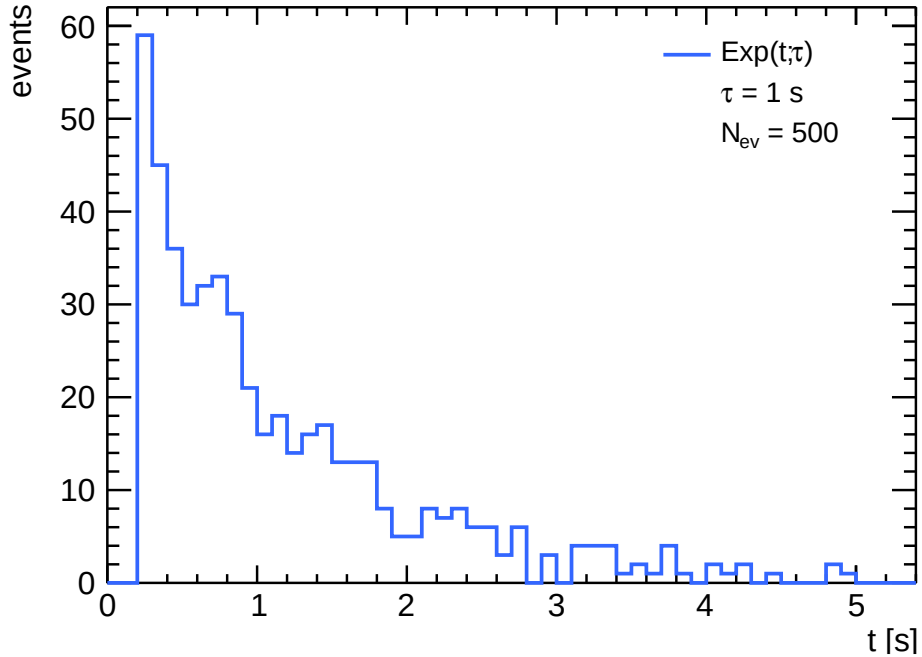
$$f_i(t;\tau) = \frac{(1/\tau)\exp(-t/\tau)}{\exp(-t_i^{\min}/\tau) - \exp(-t_i^{\max}/\tau)}$$

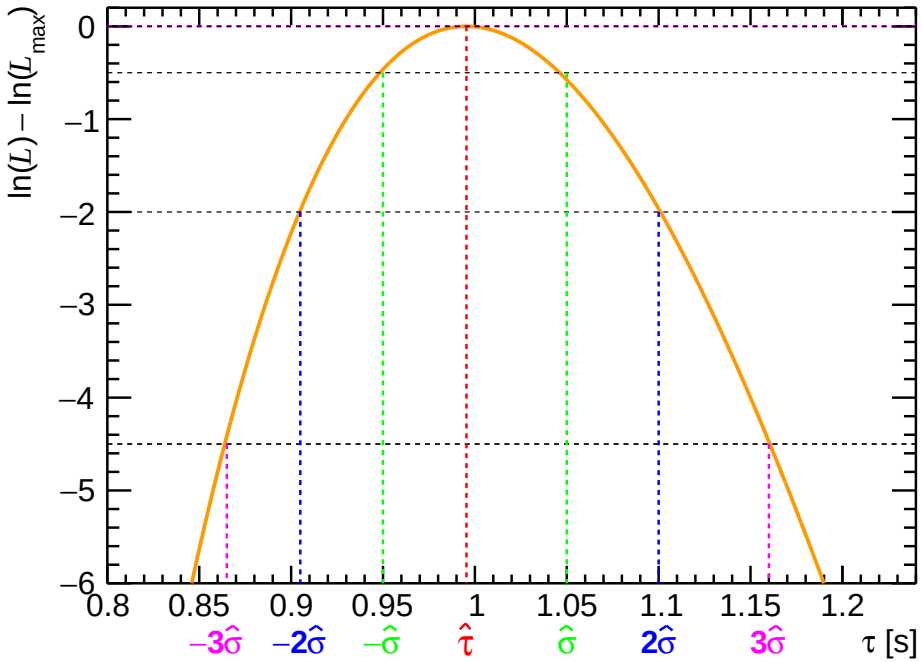
  Wyznaczamy estymator MNW:

$$\ln \mathcal{L} = -n \ln \tau - \frac{1}{\tau} \sum_{i=1}^{n} t_i - \sum_{i=1}^{n} \ln \left(\exp\left(-\frac{t_i^{\min}}{\tau} \right) - \exp\left(-\frac{t_i^{\mathrm{mx}}}{\tau} \right) \right)$$

$$\left. \frac{\partial \ln \mathcal{L}}{\partial \tau} \right|_{\tau = \hat{\tau}} = 0 \quad \Rightarrow \quad \hat{\tau} = \bar{t} - \frac{1}{n} \sum_{i=1}^{n} \frac{t_i^{\min} \exp(-t_i^{\min}/\hat{\tau}) - t_i^{\max} \exp(-t_i^{\max}/\hat{\tau})}{\exp(-t_i^{\min}/\hat{\tau}) - \exp(-t_i^{\max}/\hat{\tau})}$$

Szacowanie błędów estymatorów MNW





- Rysunki obok przedstawiaja ilustrację powyższego przykładu wykonaną metodą Monte Carlo. Przyjęto:

$$t_i^{\min} = 0.2 \text{ s} \quad \text{oraz} \quad t_i^{\max} = 5.2 \text{ s}$$

- Wartości estymatora parametru $\hat{\tau}$ oraz jego błędu (z tw. Cramera-Rao) wynoszą:

$$\hat{\tau} = 0.9954 \text{ s} \quad \text{oraz} \quad \hat{\mathcal{D}}[\hat{\tau}] = 0.0485 \text{ s}$$

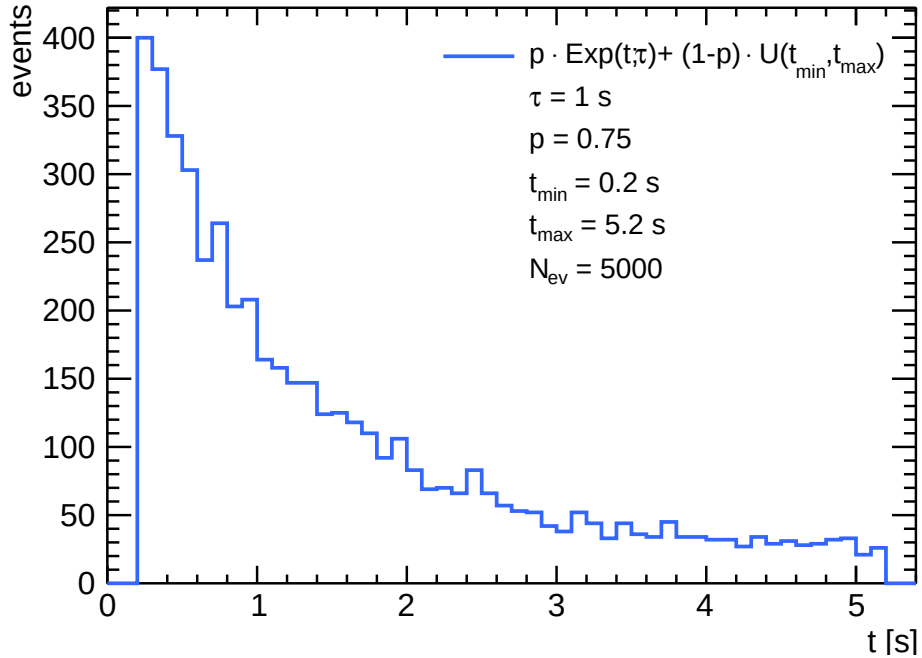
- Podobne wartości można odczytać z wykresu logarytmu funkcji wiarogodności.
- Kształt tej funkcji jest zbliżony do paraboli:

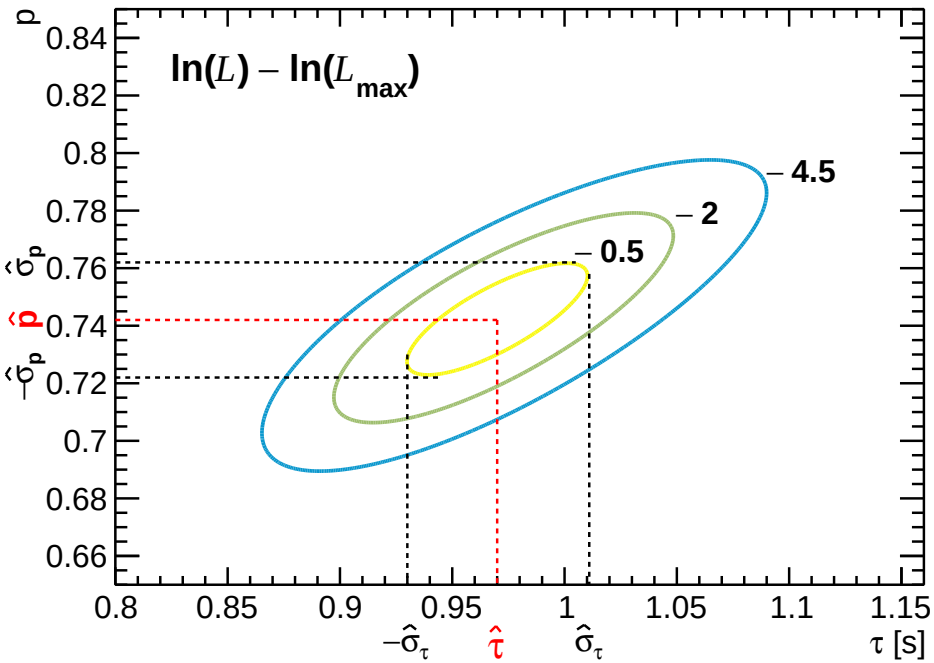
$$\ln \mathcal{L}(\theta) \approx \ln \mathcal{L}_{\max} - \frac{(\theta - \hat{\theta})^2}{2\hat{\mathcal{V}}[\hat{\theta}]}$$

a wartości $\hat{\sigma}$, $2\hat{\sigma}$ oraz $3\hat{\sigma}$ dostajemy odpowiednio dla:

$$\ln \mathcal{L}(\theta) - \ln \mathcal{L}_{\max} \approx -0.5, \quad -2.0, \quad -4.5$$

Szacowanie błędów estymatorów MNW





- Rozpatrzmy przypadek, gdy w naszym pomiarze występuje tło:

$$f(t;\tau,p) = p \cdot \mathcal{E}(t;\tau,t^{\min},t^{\max}) + \\ +(1-p)\cdot\mathcal{U}(t^{\min},t^{\max})$$

- Z wykresu konturowego wielkości $\ln\mathcal{L}(\tau,p) - \ln\mathcal{L}_{\max}$ można odczytać wartości estymatorów $\hat{\tau}$ i $\hat{p}$ oraz ich niepewności:

$$\ln\mathcal{L}(\tau,p) - \ln\mathcal{L}_{\max} \approx -\frac{1}{2(1-\rho^2)} \times \\ \left[\left(\frac{\tau-\hat{\tau}}{\sigma_{\hat{\tau}}}\right)^2 + \left(\frac{p-\hat{p}}{\sigma_{\hat{p}}}\right)^2 - 2\rho\left(\frac{\tau-\hat{\tau}}{\sigma_{\hat{\tau}}}\right)\left(\frac{p-\hat{p}}{\sigma_{\hat{p}}}\right)\right]$$

- Mierząc kąt nachylenia większej półosi elipsy do osi poziomej można wyznaczyć estymator współczynnika korelacji korzystając z formuły:

$$\operatorname{tg} 2\alpha = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}$$