Rachunek prawdopodo-bieństwa i statystyka

Wykład 10

Funkcja wiarogodności

Niech będzie dana próba losowa prosta o liczebności n z rozkładu $f(x;\theta)$.

Funkcją wiarogodności dla próby x_i nazywamy wielkość:

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

Uwaga: 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:

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

Ponieważ:

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

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

Twierdzenie Cramera-Rao

Estymator o minimalnej wariancji nazywamy efektywnym (najefektywiejszym)
Efektywnością estymatora nazywamy stosunek jego wariancji do $\mathcal{V}_{\min}$.

W przypadku n parametrów rozkładu $\vec{\theta} = (\theta_1, \ldots, \theta_n)$ nierówność Cramera-Rao ma charakter macierzowy, a minimalna wariancja wynosi:

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

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

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

<u>Uwaga:</u> Często wykorzystujemy $\mathbf{V}_{\min}$ jako przybliżenie macierzy kowariancji wstawiając w drugich pochodnych zamiast parametru jego estymatę.

Metoda największej wiarogodności

Zasada największej wiarogodności: za estymatę nieznanego parametru θ powinniśmy wybrać taką liczbę $\hat{\theta}$ dla której funkcja wiarogodności osiąga maksimum:

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

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

Często stosujemy następujące równanie do znalezienia estymaty MNW:

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

W przypadku k parametrów $\vec{\theta} = \left(\theta_1, ..., \theta_k\right)$ musimy rozwiązać układ k równań:

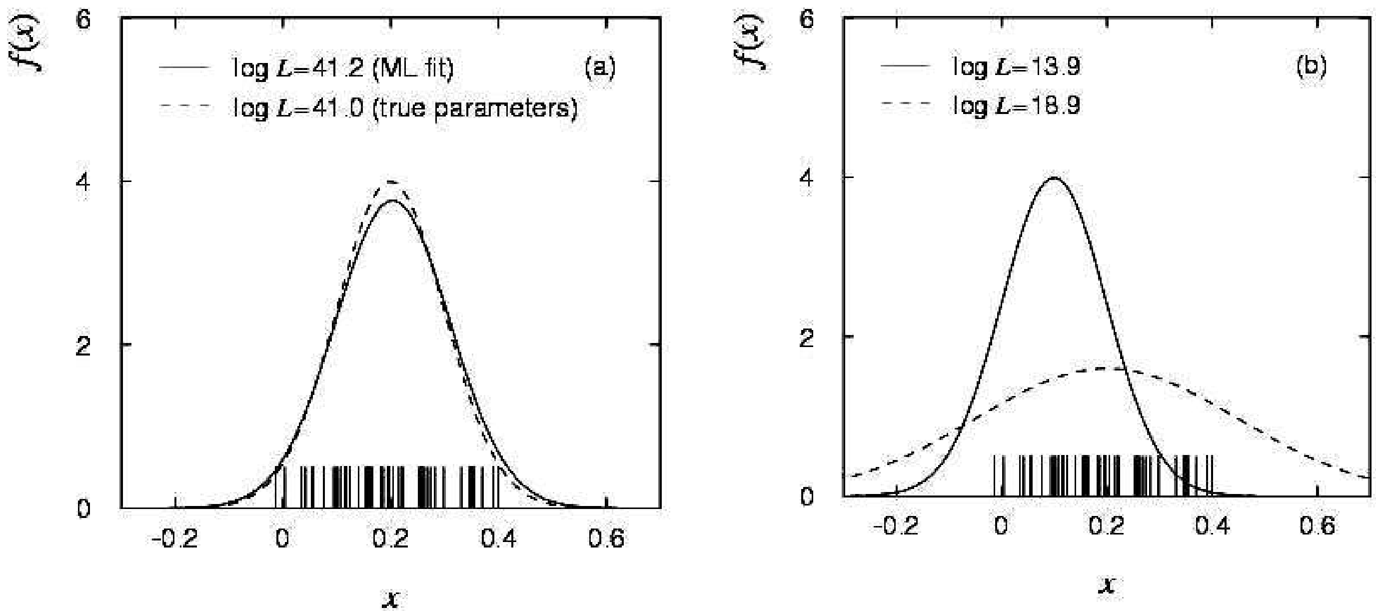
$$\frac{\partial}{\partial\theta_j}\ln\mathcal{L}(\vec{x};\vec{\theta}) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta_j}\ln f(x_i;\vec{\theta}) = 0$$

przy warunku ujemnej określoności macierzy:

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

Metoda największej wiarogodności

Jeśli oszacowana wartość parametru jest bliska wartości prawdziwej, to oczekujemy dużego p-twa otrzymania z eksperymentu danych takich jakie rzeczywiście uzyskaliśmy.



Wygenerowanych 50 przypadków z rozkładu normalnego o $\mu = 0.2$ i $\sigma = 0.1$

Wyniki otrzymane z maksymalizacji funkcji wiarogodności $\hat{\mu} = 0.204$ oraz $\hat{\sigma} = 0.106$

Odstępstwa od wartości prawdziwych są miarą błędów statystycznych.

W przypadku estymat parametrów bardzo odbiegających od wartości prawdziwych otrzymujemy znacznie mniejsze wartości funkcji wiarogodności.

Własności estymatorów MNW

Przykład: Wyznaczenie estymatora parametrów τ i λ rozkładu wykładniczego.

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

$$\frac{\partial\ln\mathcal{L}}{\partial\tau}=-\frac{n}{\tau}+\frac{1}{\tau^2}\sum_{i=1}^{n}t_i \quad\Rightarrow\quad -\frac{n}{\hat{\tau}}+\frac{1}{\hat{\tau}^2}\sum_{i=1}^{n}\mathrm{t}_i=0 \quad\Rightarrow\quad \hat{\tau}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{t}_i=\overline{\mathrm{t}}$$

Podobnie wyznaczamy estymator parametru λ:

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

$$\frac{\partial\ln\mathcal{L}}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}t_i \quad\Rightarrow\quad \frac{n}{\hat{\lambda}}-\sum_{i=1}^{n}\mathrm{t}_i=0 \quad\Rightarrow\quad \hat{\lambda}=\frac{n}{\sum_{i=1}^{n}\mathrm{t}_i}=\frac{1}{\overline{\mathrm{t}}}$$

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

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

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

Szacowanie błędów estymatorów MNW

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

<u>Przykład:</u> nieobciążony estymator parametru τ $\mathcal{E}[\hat{\tau}] = \tau$

i tylko asymptotycznie nieobciążony estymator parametru λ: $\mathcal{E}\left[\hat{\lambda}\right] = \frac{n}{n-1}\lambda$

Wariancja estymatora największej wiarogodności:

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

Wariancja przyjmuje postać funkcji ocenianego parametru. Aby uzyskać wartość liczbową musimy wstawić jego estymatę (a więc jest to przybliżenie). W przypadku k parametrów musimy znaleźć macierz kowariancji:

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

W ten sposób można wyznaczyć wariancję estymatora tylko w prostych przypadkach. Najczęściej powyższe całki są trudne/niemożliwe do rozwiązania analitycznie i musimy stosować bądź przybliżenia, bądź metodę Monte Carlo.

Szacowanie błędów estymatorów MNW

Przykład: Wariancja estymatora parametru rozkładu wykładniczego:

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

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

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

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

$$\hat{\mathcal{V}}[\hat{\theta}] = -\left(\left. \frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2} \right|_{\theta=\hat{\theta}} \right)^{-1}$$

lub w przypadku większej liczby parametrów:

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

Zachowanie asymptotyczne $\mathcal{L}(\theta)$

Przykład: Estymator wariancji estymatora parametru rozkładu wykładniczego.

$$\hat{\mathcal{V}}[\hat{\tau}] = -\left(\frac{\partial^2 \ln\mathcal{L}}{\partial\tau^2}\bigg|_{\tau=\hat{\tau}}\right)^{-1} = \frac{\hat{\tau}^2}{n} \qquad \hat{\mathcal{V}}[\hat{\lambda}] = -\left(\frac{\partial^2 \ln\mathcal{L}}{\partial\lambda^2}\bigg|_{\lambda=\hat{\lambda}}\right)^{-1} = \frac{\hat{\lambda}^2}{n}$$

Uwaga: Estymatory metody największej wiarogodności są zgodne.

Zachowanie funkcji wiarogodności jako funkcji parametru θ w okolicach wartości tego parametru określone jego estymatą MNW:

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

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

Funkcja wiarogodności jako funkcja parametru θ ma w okolicach maksimum w przybliżeniu kształt rozkładu Gaussa.

Szacowanie błędów estymatorów MNW

Metoda graficzna szacowania błędów statystycznych estymatorów MNW.

$$\ln\mathcal{L}(\theta) \approx \ln\mathcal{L}_{max} - \frac{(\theta-\hat{\theta})^2}{2\hat{\sigma}^2_{\hat{\theta}}} \quad \Rightarrow \quad \mathcal{L}\left(\hat{\theta} \pm \hat{\sigma}_{\hat{\theta}}\right) \approx \ln\mathcal{L}_{max} - \frac{1}{2}$$

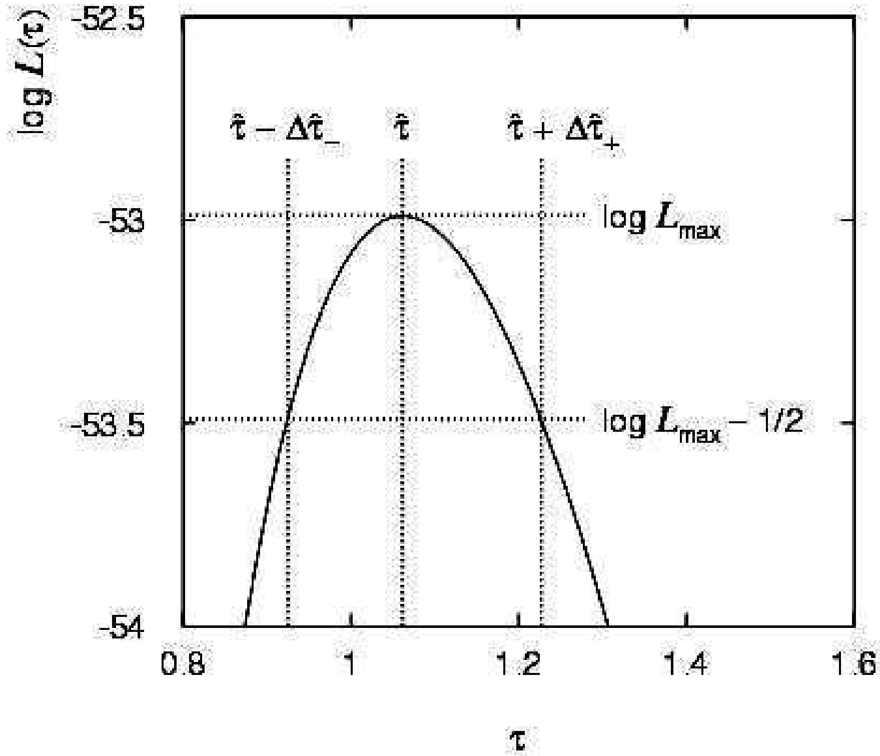
Przykład: Estymator wariancji estymatora parametru rozkładu wykładniczego.

$$\hat{\tau} = 1.062$$

$$\Delta\hat{\tau}_- = 0.137$$

$$\Delta\hat{\tau}_+ = 0.165$$

$$\hat{\sigma}_{\hat{\tau}} \approx \Delta\hat{\tau}_- \approx \Delta\hat{\tau}_+ \approx 0.15$$



Odstępstwo od paraboli ze względu na małą statystykę (n=50)

Metoda największej wiarogodności

<u>Przykład:</u> Wyznaczenie parametrów μ oraz σ rozkładu normalnego.

$$\mathcal{L}(x;\mu,\sigma)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\}$$

$$\ln\mathcal{L}=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Obliczamy pochodne funkcji wiarogodności po parametrach μ oraz σ:

$$\begin{cases}\dfrac{\partial\ln\mathcal{L}}{\partial\mu}=\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)\\[2ex] \dfrac{\partial\ln\mathcal{L}}{\partial\sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2\end{cases}\Rightarrow\begin{cases}\dfrac{1}{\hat\sigma^2}\sum_{i=1}^{n}(x_i-\hat\mu)=0 \quad\Rightarrow\quad \hat\mu=\dfrac{1}{n}\sum_{i=1}^{n}x_i\\[2ex] -\dfrac{n}{2\hat\sigma^2}+\dfrac{1}{2\hat\sigma^4}\sum_{i=1}^{n}(x_i-\hat\mu)^2=0 \quad\Rightarrow\quad \hat\sigma^2=\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\hat\mu)^2\equiv S_x^2\end{cases}$$

Estymatory wariancji estymatorów parametrów μ oraz σ:

$$\hat{\mathcal{V}}[\hat\mu]=-\left(\left.\frac{\partial^2\ln\mathcal{L}}{\partial\mu^2}\right|_{\substack{\mu=\hat\mu\\ \sigma=\hat\sigma}}\right)^{-1}=-\left(\left.-\frac{n}{\sigma^2}\right|_{\substack{\mu=\hat\mu\\ \sigma=\hat\sigma}}\right)^{-1}=\frac{\hat\sigma^2}{n}\qquad \hat{\mathcal{V}}[\hat\sigma^2]=-\left(\left.\frac{\partial^2\ln\mathcal{L}}{\partial(\sigma^2)^2}\right|_{\sigma^2=\hat\sigma^2}\right)^{-1}=\frac{2\hat\sigma^4}{n}$$

Metoda największej wiarogodności

Przykład: Opór każdego z n różnych oporników zmierzono niezależnie 2 razy. Dokładność każdego z pomiarów jest taka sama, a prawdziwe wartości oporu nie są znane. Przyjmując, że każdy z pomiarów opisany jest rozkładem Gaussa o tej samej dyspersji σ, ale różnych wartościach oczekiwanych μ_i (i = 1, ..., n), odpowiadających różnym oporom, znajdź estymatory wielkości μ_i i σ.

$$\mathcal{L}(x,y;\vec{\mu},\sigma)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x_i-\mu_i)^2}{2\sigma^2}\right)\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(y_j-\mu_j)^2}{2\sigma^2}\right)=$$

$$=\frac{1}{\left(\sqrt{2\pi}\sigma\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_i)^2\right)\frac{1}{\left(\sqrt{2\pi}\sigma\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(y_j-\mu_j)^2\right)$$

$$\ln\mathcal{L}=-\frac{1}{2}\ln 2\pi-\frac{1}{2}n\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_i)^2-\frac{1}{2}\ln 2\pi-\frac{1}{2}n\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu_i)^2$$

$$\frac{\partial\ln\mathcal{L}}{\partial\mu_i}=\frac{x_i-\mu_i}{\sigma^2}+\frac{y_i-\mu_i}{\sigma^2}\quad\Rightarrow\quad\frac{x_i-\hat{\mu}_i}{\hat{\sigma}^2}+\frac{y_i-\hat{\mu}_i}{\hat{\sigma}^2}=0\quad\Rightarrow\quad\boxed{\hat{\mu}_i=\frac{1}{2}(x_i+y_i)}$$

$$\frac{\partial\ln\mathcal{L}}{\partial\sigma^2}=-\frac{1}{2}n\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu_i)^2-\frac{1}{2}n\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i-\mu_i)^2$$

$$-\frac{n}{\hat{\sigma}^2}+\frac{1}{2\hat{\sigma}^4}\sum_{i=1}^{n}\left[(x_i-\hat{\mu}_i)^2+(y_i-\hat{\mu}_i)^2\right]=0\quad\Rightarrow\quad\hat{\sigma}^2=\boxed{\frac{1}{4n}\sum_{i=1}^{n}(x_i-y_i)^2}$$

Metoda największej wiarogodności

<u>Przykład:</u> Wyznaczenie estymatora parametru rozkładu Poissona.

$$\mathcal{L}(k;\mu)=\prod_{i=1}^{n}\frac{\mu^{k_i}}{k_i\,!}e^{-\mu} \quad\Rightarrow\quad \ln\mathcal{L}=\sum_{i=1}^{n}\left(\ln\frac{\mu^{k_i}}{k_i\,!}-\mu\right)=\sum_{i=1}^{n}k_i\ln\mu-\sum_{i=1}^{n}\ln k_i\,!-n\mu$$

$$\frac{\partial\ln\mathcal{L}}{\partial\mu}=\frac{1}{\mu}\sum_{i=1}^{n}k_i-n \quad\Rightarrow\quad \frac{1}{\hat{\mu}}\sum_{i=1}^{n}\mathrm{k}_i-n=0 \quad\Rightarrow\quad \hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{k}_i$$

Wartość oczekiwana i wariancja:

$$\mathcal{E}[\hat{\mu}]=\left\langle\frac{1}{n}\sum_{i=1}^{n}\mathrm{k}_i\right\rangle=\frac{1}{n}\sum_{i=1}^{n}\langle\mathrm{k}_i\rangle=\frac{1}{n}n\mu=\mu$$

$$\mathcal{V}[\hat{\mu}]=\sum_{k=0}^{\infty}\left[(\hat{\mu}-\mu)^2 L(k;\mu)\right]=\sum_{k=0}^{\infty}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\mathrm{k}_i-\mu\right)^2\prod_{l=1}^{n}\frac{\mu^{k_i}}{k_l\,!}e^{-\mu}\right]=\sum_{k=0}^{\infty}\left[\left(\frac{1}{n}\sum_{i=1}^{n}(\mathrm{k}_i-\mu)\right)^2\prod_{l=1}^{n}\frac{\mu^{k_i}}{k_l\,!}e^{-\mu}\right]=$$

$$=\sum_{k=0}^{\infty}\left[\frac{1}{n^2}\left(\sum_{i=1}^{n}(\mathrm{k}_i-\mu)^2+\sum_{i\neq j=1}^{n}(\mathrm{k}_i-\mu)(\mathrm{k}_j-\mu)\right)\prod_{l=1}^{n}\frac{\mu^{k_i}}{k_l\,!}e^{-\mu}\right]=$$

$$=\frac{1}{n^2}\sum_{k=0}^{\infty}\left(\sum_{i=1}^{n}(k_i-\mu)^2\prod_{l=1}^{n}\frac{\mu^{k_i}}{k_l\,!}e^{-\mu}\right)=\frac{1}{n^2}n\mathcal{V}[\mathrm{k}]=\frac{\mu}{n}$$

Średnia ważona

Dana jest próbka prosta złożona z n elementów x_i (i=1,...,n), każdy z rozkładu normalnego o różnej i znanej dyspersji σ_i, ale nieznanej i identycznej wartości oczekiwanej μ. (np. pomiary danej wielkości za pomocą różnych przyrządów).

$$\mathcal{L}(\vec{x};\mu,\sigma_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma_i^2} \right)$$

$$\ln \mathcal{L} = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{n} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma_i^2} \quad \Rightarrow \quad \frac{\partial \ln \mathcal{L}}{\partial \mu} = \sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma_i^2}$$

$$\sum_{i=1}^{n} \frac{(x_i-\hat{\mu})}{\sigma_i^2} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} = \hat{\mu}\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \quad \Rightarrow \quad \hat{\mu} = \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \Big/ \sum_{i=1}^{n} \frac{1}{\sigma_i^2}$$

Średnia ważona: $\hat{\mu} = \sum_{i=1}^{n} w_i x_i$ gdzie $w_i = \dfrac{\frac{1}{\sigma_i^2}}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}}$

Średnia ważona

Wartość oczekiwana i wariancja:

$$\int_{-\infty}^{\infty} \exp\left(-ax^2\right) d\,x = \sqrt{\frac{\pi}{a}}$$

$$\int_{-\infty}^{\infty} x^2 \exp\left(-ax^2\right) d\,x = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

$$\mathcal{E}[\hat{\mu}] = \sum_{i=1}^{n} w_i \langle x_i \rangle = \mu \sum_{i=1}^{n} w_i = \mu$$

$$\mathcal{V}[\hat{\mu}] = \int_{-\infty}^{\infty} (\hat{\mu}-\mu)^2 L(x;\mu,\sigma) dx = \int_{-\infty}^{\infty} \left(\sum_{i=1}^{n} w_i x_i - \mu\right)^2 \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j-\mu)^2}{2\sigma_j^2}\right) dx =$$

$$= \int_{-\infty}^{\infty} \left(\sum_{i=1}^{n} w_i (x_i - \mu)\right)^2 \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j-\mu)^2}{2\sigma_j^2}\right) dx =$$

$$= \int_{-\infty}^{\infty} \left(\sum_{i=1}^{n} w_i^2 (x_i - \mu)^2 + \sum_{i\neq k=1}^{n} w_i w_k (x_i-\mu)(x_k-\mu)\right) \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j-\mu)^2}{2\sigma_j^2}\right) dx =$$

$$= \sum_{i=1}^{n} w_i^2 \sigma_i^2 \prod_{\substack{j=1 \\ j\neq i}}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \sqrt{2\pi}\sigma_j = \sum_{i=1}^{n} w_i^2 \sigma_i^2 = \sum_{i=1}^{n} \left(\frac{\frac{1}{\sigma_i^2}\sigma_i}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}}\right)^2 = \frac{1}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}}$$

Metoda największej wiarogodności

<u>Przykład:</u> Prędkość dźwięku mierzona w powietrzu dwiema różnymi metodami wynosi: $v_1=340 \pm 9$ [m/s] oraz $v_2=350 \pm 18$ [m/s]. Znajdź najlepszą estymatę prędkości dźwięku.

Obliczamy wagi: $w_1 = \frac{1}{9^2} = 0.013$ $\qquad$ $w_2 = \frac{1}{18^2} = 0.003$

Średnia ważona prędkość dźwięku oraz jej błąd:

$$v_w = \frac{w_1 v_1 + w_2 v_2}{w_1 + w_2} = 342 \ \ [\text{m/s}] \qquad \sigma_w = \sqrt{(w_1 + w_2)^{-1}} = 8 \ \ [\text{m/s}]$$