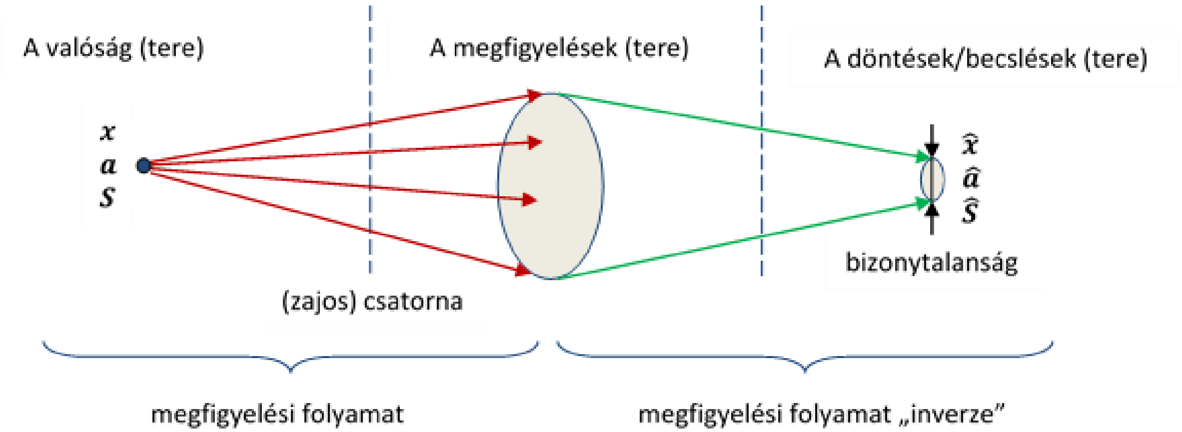
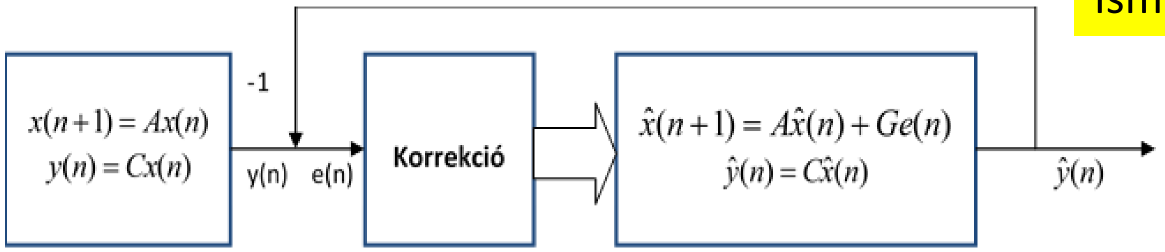
Méréselmélet

3. fejezet: A becsléselmélet alapjai

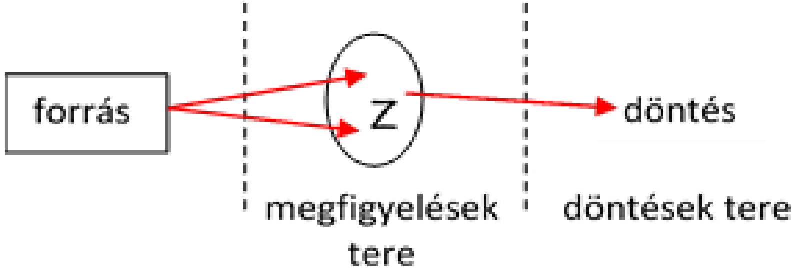
2022. március 2.

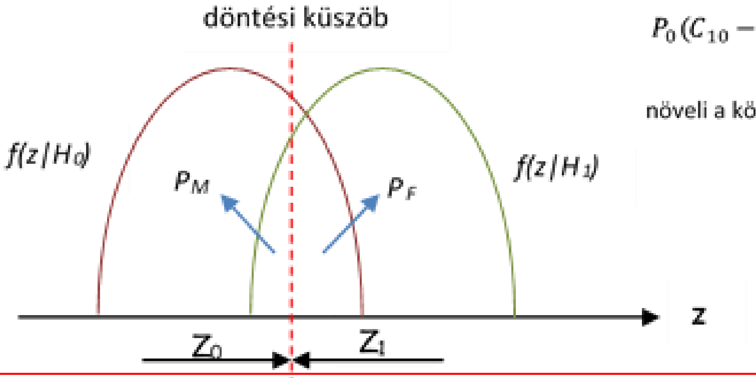


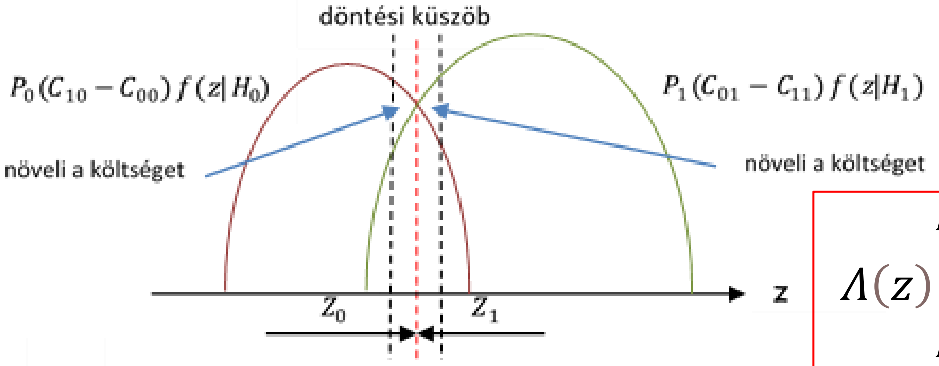


$$P_0(C_{10} - C_{00}) f(z|H_0) = P_1(C_{01} - C_{11}) f(z|H_1).$$

2. A döntéselmélet alapjai







$$\Lambda(z) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

$$R = C_{00}P_0P(H_0|H_0) \quad + \quad C_{10}P_0P(H_1|H_0) \quad + \quad C_{01}P_1P(H_0|H_1) \quad + \quad C_{11}P_1P(H_1|H_1)$$

$$R = C_{10}P_0 + C_{11}P_1 + P_0(C_{00} - C_{10}) \int_{Z_0} f(z|H_0)dz + P_1(C_{01} - C_{11}) \int_{Z_0} f(z|H_1)dz.$$

$$\Lambda(z) = \left.\frac{f(z|H_1)}{f(z|H_0)}\right|_{z_{küszöb}} = $$

$$= \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} = \eta$$

1. Példa: jelen van-e a jel a megfigyelésekben vagy nincsen?

$$\Lambda(z) = \prod_{k=0}^{N-1} \frac{f(z_k|H_1)}{f(z_k|H_0)} = \prod_{k=0}^{N-1} \frac{e^{-\frac{(z_k-a)^2}{2\sigma_n^2}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta,$$

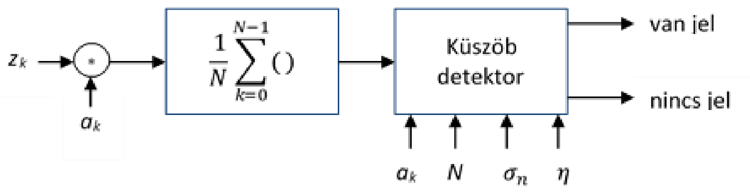
$$\frac{1}{N}\sum_{k=0}^{N-1} z_k \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma_n^2}{Na} \ln \eta + \frac{a}{2},$$

2. Példa: Változó amplitúdójú jel detektálása: jelen van-e a jel a megfigyelésekben vagy nincsen?

Ismétlés

$$\Lambda(z) = \prod_{k=0}^{N-1} \frac{f(z_k|H_1)}{f(z_k|H_0)} = \prod_{k=0}^{N-1} \frac{e^{-\frac{(z_k-a_k)^2}{2\sigma_n^2}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta,$$

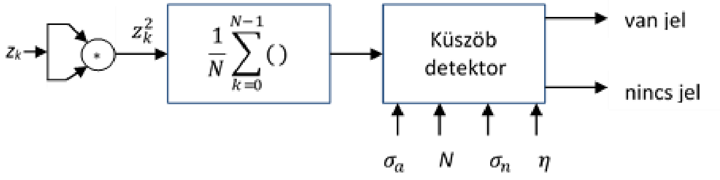
$$\frac{1}{N}\sum_{k=0}^{N-1} a_k z_k \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma_n^2}{N} \ln\eta + \frac{1}{2N}\sum_{k=0}^{N-1} a_k^2,$$

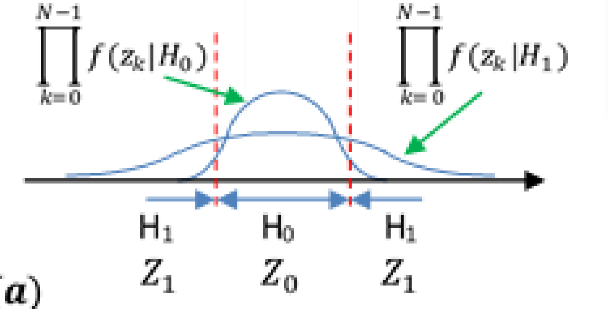


3. példa: Véletlen amplitúdójú jel detektálása: jelen van-e a jel a megfigyelésekben vagy nincsen?

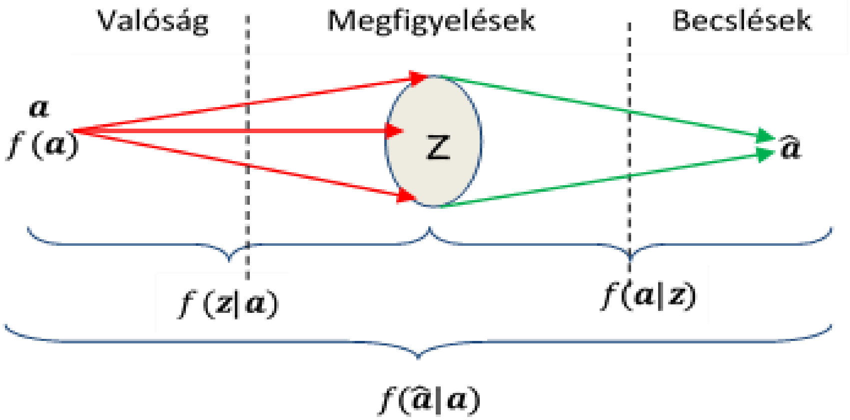
$$\Lambda(z) = \prod_{k=0}^{N-1} \frac{f(z_k|H_1)}{f(z_k|H_0)} = \prod_{k=0}^{N-1} \left[\frac{\sigma_n}{\sqrt{\sigma_a^2+\sigma_n^2}} \frac{e^{-\frac{z_k^2}{2(\sigma_a^2+\sigma_n^2)}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}} \right] \underset{H_0}{\overset{H_1}{\gtrless}} \eta.$$

$$\frac{1}{N}\sum_{k=0}^{N-1} z_k^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{2\sigma_n^2(\sigma_a^2+\sigma_n^2)}{\sigma_a^2}\left[\frac{1}{2}\ln\frac{\sigma_a^2+\sigma_n^2}{\sigma_n^2} + \frac{1}{N}\ln\eta\right]$$

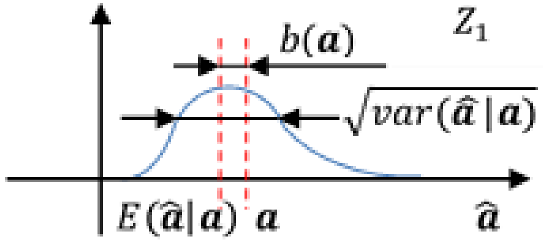




3. A becsléselmélet alapjai

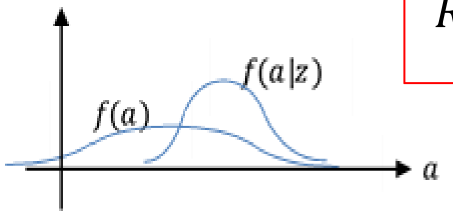


A cél az $\boldsymbol{a}$ paraméter(vektor) $\hat{\boldsymbol{a}}$ becslőjének meghatározása az ábrán látható előzetesen ismert (és az azokból származtatott) sűrűségfüggvények ismeretében.



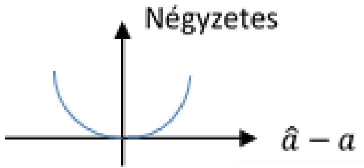
3.1. Bayes becslők

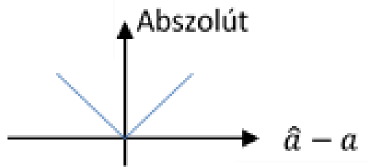
$$f(a|z) = \frac{f(z|a)f(a)}{f(z)},$$

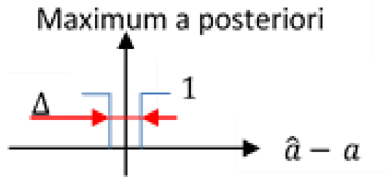


$$R(\hat{a},a) = E\{C(\hat{a},a)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(\hat{a},a) f(a,z)\,da\,dz$$

Itt $C(\hat{a},a)$ az ún. **költségfüggvény**, vagy **kockázatfüggvény**, …







Ismétlés

3.1.1. Minimális átlagos négyzetes hibájú becslés

$$\hat{a}_{MS} = \int_{-\infty}^{\infty} a f(a|z) da$$

A legjobb becslés az *a posteriori* **várható érték**!

3.1.2. Minimális átlagos abszolút hibájú becslés

$$\int_{-\infty}^{\hat{a}_{ABS}} f(a|z)da = \int_{\hat{a}_{ABS}}^{\infty} f(a|z)da$$

$\hat{a}_{ABS} = f(a|z)$ **mediánja.**

3.1.3. Maximum a posteriori (MAP) becslés

$\hat{a}_{MAP} = f(a|z)$ **maximumhelye.**

3.1.4. Bayes becslő Gauss eloszlások esetén

Ha “minden” Gauss eloszlású, akkor az a posteriori sűrűségfüggvény momentumai **explicit formában** megadhatók.

Tegyük fel, hogy az **additív zajjal** terhelt megfigyelés az alábbi összefüggéssel írható le: $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{n}$

A megfigyelés modellje lineáris!

Itt $\dim \boldsymbol{a} = p, \dim \boldsymbol{z} = q, \dim \boldsymbol{U} = q * p$. Az $\boldsymbol{U}$ az ún. **megfigyelési mátrix**.

Az *a posteriori* várható érték explicit formája ilyen körülmények között:

$$\hat{\boldsymbol{a}}_{MS} = \boldsymbol{\mu}_{\boldsymbol{a}|\boldsymbol{z}} = \underbrace{\boldsymbol{\mu}_{\boldsymbol{a}}}_{\text{apriori ismeret}} + \underbrace{[\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U} + \boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}(\boldsymbol{z} - \boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}})}_{\text{korrekció } \boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}} \text{ függvényében}}$$

Az *a posteriori* kovariancia explicit formája ilyen körülmények között:

$$cov\{a, a|\boldsymbol{z}\} = \boldsymbol{\Sigma}_{aa|\boldsymbol{z}} = [\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U} + \boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}$$

Megjegyzés: $\hat{\boldsymbol{a}}_{MS} = \hat{\boldsymbol{a}}_{ABS} = \hat{\boldsymbol{a}}_{MAP}$, mert az a Gauss *a posteriori* sűrűségfüggvény szimmetrikus.

1. Példa: Mérendő egy ellenállás értékét úgy, hogy ismert áram által ejtett feszültséget mérünk. *N* megfigyelést végzünk.

“$U = IR + zaj$” a modellünk. A megfigyelt értékek: $z_k = a + n_k, k = 0, 1, \ldots, N-1,$

a az ismeretlen paraméter (ellenállás), n_k az additív zaj mintája.

Tegyük fel, hogy az **ellenállás értéke** és a **megfigyelési zaj** egyaránt **Gauss eloszlású** valószínűségi változóknak tekinthetők.

A zaj megfigyelési értékei **korrelálatlanok**. Tegyük fel, hogy **ismert** μ_a és σ_a^2 $\rightarrow$ az ismeretlen paraméter jellemzése.

a zaj várható értéke $\mu_n = 0$, $cov\{n_k, n_j\} = \sigma_n^2 \delta_{kj}$, ahol $\delta_{kj} = \begin{cases} 1 \text{ ha } k = j \\ 0 \text{ ha } k \neq j \end{cases}$ $\rightarrow$ a csatorna jellemzése.

Vektoros alakban: $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{n}$ $\quad \boldsymbol{z}^T = [z_0, z_1, \ldots, z_{N-1}]$, $\boldsymbol{n}^T = [n_0, n_1, \ldots, n_{N-1}]$, $\boldsymbol{U} = [1 \quad 1 \quad \cdots \quad 1]^T$, $\boldsymbol{\Sigma}_{\boldsymbol{nn}} = \sigma_n^2 \boldsymbol{I}$

Ismétlés

$$\hat{\boldsymbol{a}}_{MS}=\boldsymbol{\mu}_{\boldsymbol{a}|\boldsymbol{z}}=\underbrace{\boldsymbol{\mu}_{\boldsymbol{a}}}_{\text{apriori ismeret}}+\underbrace{[\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U}+\boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}})}_{\text{korrekció } \boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}} \text{ függvényében}}$$

$$cov\{a,a|\boldsymbol{z}\}=\boldsymbol{\Sigma}_{aa|\boldsymbol{z}}=[\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U}+\boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}$$

$$\hat{a}_{MS}=\mu_{a|z}=\mu_a+\left[\frac{N}{\sigma_n^2}+\frac{1}{\sigma_a^2}\right]^{-1}\left(\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1}z_k-\frac{N}{\sigma_n^2}\mu_a\right)=\underbrace{\mu_a}_{\text{előzetes ismeret}}+\frac{N\frac{\sigma_a^2}{\sigma_n^2}}{1+N\frac{\sigma_a^2}{\sigma_n^2}}\Big(\underbrace{\frac{1}{N}\sum_{k=0}^{N-1}z_k}_{\text{minták átlaga}}-\mu_a\Big)=\frac{\mu_a+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1}z_k}{1+N\frac{\sigma_a^2}{\sigma_n^2}}$$

Ha $\sigma_a << \sigma_n$, akkor $\hat{a}_{MS}\cong\mu_a$,
ha $\sigma_a >> \sigma_n$vagy $N\to\infty$, akkor $\hat{a}_{MS}\cong\frac{1}{N}\sum_{k=0}^{N-1}z_k$.

predikció ➕ korrekció a méréssel szerzett információ alapján.

A becslési hiba **varianciája**: az *a posteriori* kovariancia:

$$cov\{a,a|\boldsymbol{z}\}=\boldsymbol{\Sigma}_{aa|\boldsymbol{z}}=[\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U}+\boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}=var\{\tilde{a}\}=\frac{\sigma_a^2}{1+N\frac{\sigma_a^2}{\sigma_n^2}}$$

$\boxed{\tilde{a}=\hat{a}-a}$

Ha $\sigma_a << \sigma_n$, akkor $var\{\tilde{a}\}\cong\sigma_a^2$,
ha $\sigma_a >> \sigma_n$vagy $N\to\infty$, akkor $var\{\tilde{a}\}\cong\frac{1}{N}\sigma_n^2$.

A becslés feltételesen torzított: $b(a)=E\{\hat{a}_{MS}|a\}-a=$

$$=E\left\{\frac{\mu_a+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1}z_k}{1+N\frac{\sigma_a^2}{\sigma_n^2}}\right\}-a=\frac{\mu_a-a}{1+N\frac{\sigma_a^2}{\sigma_n^2}}.$$

Itt felhasználtuk, hogy $E\{z_k\}=a$.

2. Példa: Mérendő egy ismert jel ismeretlen amplitúdója.

$z_k=as_k+n_k,\ k\ =\ 0,1,\dots,N-1.$ Az ismeretlen a amplitúdó és n_k Gauss eloszlású. $E\{a\}=\mu_a$, $var\{a\}=\sigma_a^2$,
$E\{n_k\}=0$, $cov\{n_i,n_j\}=\sigma_n^2\delta_{ij}$,$cov\{a,n_i\}=0$ $\forall i$ -re, $\forall j$ -re.

Most használjuk a **maximum a posteriori** (MAP) becslést!

$$\boxed{\left.\frac{\partial f(a|\boldsymbol{z})}{\partial a}\right|_{a=\hat{a}_{MAP}}=0}\ \to\ \left.\frac{\partial\ln f(\boldsymbol{z}|a)}{\partial a}+\frac{\partial\ln f(a)}{\partial a}\right|_{a=\hat{a}_{MAP}}=0$$

Most $f(a)=\frac{1}{\sqrt{2\pi}\sigma_a}e^{-\frac{(a-\mu_a)^2}{2\sigma_a^2}}\ \to\ \frac{\partial\ln f(a)}{\partial a}=-\frac{a-\mu_a}{\sigma_a^2}$

$$\boxed{f(a|z)=\frac{f(z|a)f(a)}{f(z)}}\quad\boxed{\frac{\partial\ln f(z)}{\partial a}=0}\quad f(\boldsymbol{z}|a)=\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}e^{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1}(z_k-as_k)^2}\to\ \frac{\partial\ln f(\boldsymbol{z}|a)}{\partial a}=\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1}s_k(z_k-as_k)$$

$$\boxed{\left.\frac{\partial \ln f(\mathbf{z}|a)}{\partial a}+\frac{\partial \ln f(a)}{\partial a}\right|_{a=\hat{a}_{MAP}}=0} \qquad \frac{\partial \ln f(a)}{\partial a}=-\frac{a-\mu_a}{\sigma_a^2} \qquad \frac{\partial \ln f(\mathbf{z}|a)}{\partial a}=\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} s_k\left(z_k-as_k\right)$$

$$\left.\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} s_k\left(z_k-as_k\right)-\frac{a-\mu_a}{\sigma_a^2}\right|_{a=\hat{a}_{MAP}}=0 \;\rightarrow\; \boxed{\hat{a}_{MAP}=\frac{\mu_a+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k z_k}{1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}} \qquad \boxed{var\{\tilde{a}\}=\frac{\sigma_a^2}{1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}}$$

Ha s_k=1, $\forall k$, akkor megkapjuk az előző példa eredményét!
Itt is azonosítható a döntéselméleti rész 2. feladatában említett „illesztett” szűrő. Természetesen $\hat{a}_{MAP}=\hat{a}_{MS}$.

Megjegyzés: A variancia kifejezésének bizonyítása:

$$var\{\tilde{a}\}=E\{(\hat{a}-a)^2\}=E\left\{\left(\frac{\mu_a+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k z_k-a-a\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}{1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}\right)^2\right\}=\frac{E\left\{\left((\mu_a-a)+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k(z_k-as_k)\right)^2\right\}}{\left(1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2\right)^2}$$

Itt $E\{(\mu_a-a)^2\}=\sigma_a^2$

$E\{(\mu_a-a)(z_k-as_k)\}=0$

mert $E\{n_k\}=0$, és

$E\{an_k\}=0, \forall k$

A második tag négyzete:

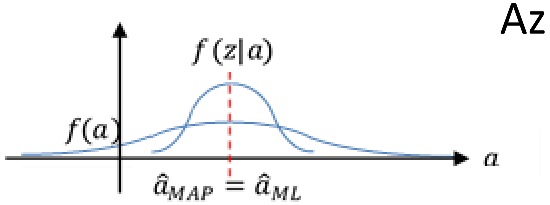
$$E\left\{\left(\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k(z_k-as_k)\right)^2\right\}=\frac{\sigma_a^4}{\sigma_n^4}\sigma_n^2\sum_{k=0}^{N-1} s_k^2 \text{ mert } E\{n_k n_j\}=0, k\neq j.$$

Ezzel:

$$var\{\tilde{a}\}=E\{(\hat{a}-a)^2\}=\frac{\sigma_a^2+\frac{\sigma_a^4}{\sigma_n^4}\sigma_n^2\sum_{k=0}^{N-1} s_k^2}{\left(1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2\right)^2}=\frac{\sigma_a^2}{1+\frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}$$

3.2. Maximum Likelihood (ML) becslő

Kiindulásunk, hogy **nem ismerjük** a mérendő mennyiség a priori valószínűségsűrűség függvényét. Ilyenkor azt feltételezzük, hogy ez a függvény „szélesen elterülő”, és ebből adódóan az *a posteriori* sűrűségfüggvény megegyezik a csatorna-karakterisztikával. Az optimális becslőt a csatorna-karakterisztika maximumához rendeljük:



$$\boxed{\left.\frac{\partial f(\mathbf{z}|\mathbf{a})}{\partial \mathbf{a}}\right|_{a=\hat{a}_{ML}}=0,} \quad \text{ill.} \quad \boxed{\left.\frac{\partial \ln f(\mathbf{z}|\mathbf{a})}{\partial \mathbf{a}}\right|_{a=\hat{a}_{ML}}=0}$$

3.2.1. Gauss-Markov (GM) becslő: Speciális ML becslő: a megfigyelési zaj **Gauss** eloszlású, a megfigyelési egyenlet **lineár** Ismétlés

N megfigyelést végzünk, $\boldsymbol{n}$ az N dimenziós zaj-vektor: $E\{\boldsymbol{n}\} = 0$, $cov\{\boldsymbol{n},\boldsymbol{n}\} = \boldsymbol{\Sigma_{nn}}$

$$\boxed{f(\boldsymbol{n}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\boldsymbol{\Sigma_{nn}}|^{\frac{1}{2}}} e^{-\frac{1}{2}\boldsymbol{n}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{n}}}$$

A megfigyelési egyenlet: $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{n}$, amellyel a csatorna karakterisztika:

$$\boxed{f(\boldsymbol{z}|\boldsymbol{a}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\boldsymbol{\Sigma_{nn}}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T\boldsymbol{\Sigma_{nn}^{-1}}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})}}$$

$$\frac{\partial \ln f(\boldsymbol{z}|\boldsymbol{a})}{\partial \boldsymbol{a}} = \frac{\partial}{\partial \boldsymbol{a}}\left[(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T\boldsymbol{\Sigma_{nn}^{-1}}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})\right]\Big|_{\boldsymbol{a}=\widehat{\boldsymbol{a}}_{GM}} = 0.$$

ahol $|\boldsymbol{\Sigma_{nn}}|$ a $\boldsymbol{\Sigma_{nn}}$ mátrix determinánsát jelöli.

$$\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})\Big|_{\boldsymbol{a}=\widehat{\boldsymbol{a}}_{GM}} = 0 \quad \rightarrow \quad \boxed{\widehat{\boldsymbol{a}}_{GM} = [\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{z}}$$

ha $[\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}$ létezik.

A becslés minőségjellemzője:

$$cov(\widehat{\boldsymbol{a}}_{GM},\widehat{\boldsymbol{a}}_{GM}) = E\{(\widehat{\boldsymbol{a}}_{GM}-\boldsymbol{a})(\widehat{\boldsymbol{a}}_{GM}-\boldsymbol{a})^T\} = E\{[\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}\underbrace{\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\overbrace{(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T}^{\text{Várható értéke: } \boldsymbol{\Sigma_{nn}}}\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}[\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}}_{\boldsymbol{I}}\} = [\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}$$

$$\boxed{cov(\widehat{\boldsymbol{a}}_{GM},\widehat{\boldsymbol{a}}_{GM}) = [\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U}]^{-1}}$$

Megjegyzések:

$$\boxed{\widehat{\boldsymbol{a}}_{MS} = \boldsymbol{\mu}_{\boldsymbol{a}|\boldsymbol{z}} = \underbrace{\boldsymbol{\mu_a}}_{\text{apriori ismeret}} + \underbrace{[\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}\boldsymbol{U} + \boldsymbol{\Sigma_{aa}^{-1}}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma_{nn}^{-1}}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu_a})}_{\text{korrekció } \boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu_a} \text{ függvényében}}}$$

Ha $\boldsymbol{\Sigma_{aa}^{-1}} = \boldsymbol{0}$, vagyis a szórás végtelen, akkor

$$\boxed{\widehat{\boldsymbol{a}}_{MS} = \widehat{\boldsymbol{a}}_{GM}}$$

A Gauss-Markov becslő torzítatlan: $E\{\widehat{\boldsymbol{a}}_{GM}\} = \boldsymbol{a}$

Példa: A megfigyelési egyenlet: $z_k = a + n_k$, $k = 0,1,\dots,N-1$, a megfigyelések függetlenek. $E\{n_k\} = 0$; $cov\{n_k,n_j\} = \sigma_n^2\delta_{kj}$

A csatorna karakterisztika:

$$\boxed{f(\boldsymbol{z}|a) = \frac{1}{(2\pi\sigma_n^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1}(z_k-a)^2},} \quad \boxed{\frac{\partial \ln f(\boldsymbol{z}|a)}{\partial a}\Big|_{\hat{a}=\hat{a}_{ML}=\hat{a}_{GM}} = \frac{N}{\sigma_n^2}\left[\frac{1}{N}\sum_{k=0}^{N-1} z_k - a\right] = 0,} \quad \boxed{\hat{a}_{ML} = \hat{a}_{GM} = \frac{1}{N}\sum_{k=0}^{N-1} z_k,}$$

azaz lineáris megfigyelési egyenlet és Gauss eloszlású csatorna zaj esetén a legjobb (Gauss-Markov) becslő **az egyszerű átlagolás!**

3.3. Becslők determinisztikus modellel jellemzett paraméterek esetén

A becslési probléma: adott $\boldsymbol{z} = \{z_k\}, k = 0,1,\dots,N-1$, azaz N mért érték, amelyek az ismeretlen $\boldsymbol{a}$ paraméter függvényei. Keressük $\boldsymbol{a}$ becslőjét $\hat{\boldsymbol{a}} = \boldsymbol{g}(z_0, z_1, \dots, z_{N-1})$, ahol $\boldsymbol{g}$ egy függvény. Az első lépés a „csatorna-karakterisztika” meghatározása: azaz a megfigyelt adatok valószínűség sűrűségfüggvényének megtalálása az $\boldsymbol{a}$ paraméter függvényében, amit $f(\boldsymbol{z}; \boldsymbol{a})$ jelöl.

Példa: Additív, **Gauss eloszlású fehér** zajjal terhelt **DC szint** mérése **egyetlen** mért adatra alapozva: $\boxed{z_0 = a + w_0}$
ahol a w_0 valószínűség sűrűségfüggvénye $\mathcal{N}(0, \sigma^2)$. (**Normális** (Gauss) eloszlás, nulla várható érték, σ^2 variancia.)

Ebben az esetben a „**csatorna-karakterisztika**”, azaz a z_0 sűrűségfüggvénye:

$$\boxed{f(z_0; a) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(z_0 - a)^2\right]}$$

(Ez már nem feltételes, $\boldsymbol{a}$ a függvény egy paramétere!)

Példa: Additív, **Gauss eloszlású fehér** zajjal terhelt **lineáris sorozat** mérése **mérési sorozatra** alapozva: $\boxed{z_k = A + Bk + w_k,}$
$k = 0,1,\dots,N-1$, w_k korrelálatlan minden mintával, és valószínűség sűrűségfüggvénye $\mathcal{N}(0, \sigma^2)$.
Legyen $\boldsymbol{a} = [A\ B]$, $\boldsymbol{z} = [z_0, z_1, \dots, z_{N-1}]$. Ebben az esetben a „**csatorna-karakterisztika**”, azaz $\boldsymbol{z}$ sűrűségfüggvénye:

$$\boxed{f(\boldsymbol{z}; \boldsymbol{a}) = \prod_{k=0}^{N-1} f(z_k; \boldsymbol{a}) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^N} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}\left(z_k - A - Bk\right)^2\right],}$$

3.3.1. Becslések minősítése

Tekintsünk egy A szintű DC mérést korrelálatlan zajban:

$z_k = A + w_k, k = 0,1,\dots,N-1.$

Tekintsük a következő becslőket: $\hat{A}_1 = \frac{1}{N}\sum_{k=0}^{N-1} z_k$, $\hat{A}_2 = z_0$. $\boxed{\text{Tegyük fel, hogy } A = 1,}$ $\boxed{\hat{A}_1 = 0.95,}$ $\boxed{\hat{A}_2 = 0.98.}$ $\boxed{\text{Melyik becslés jobb?}}$

Mivel a becslő valószínűségi változó, ezért viselkedése/minősítése a valószínűség sűrűségfüggvényével adható meg. Ha az ismert!
(Ha nem ismert, akkor statisztikai eszközökkel vagy például Monte-Carlo szimulációval.)

Torzítatlan becslő: (Nagyon szeretjük!) Az a becslő, amelyik várható értékben a helyes értéket adja: $\boxed{E(a - \hat{a}) = 0}$

$$\boxed{E(\hat{A}_1) = \frac{1}{N}\sum_{k=0}^{N-1} E(z_k) = \frac{1}{N}\sum_{k=0}^{N-1} E(A + w_k) = \frac{1}{N}\sum_{k=0}^{N-1}(A + 0) = A,}$$

$$\boxed{E(\hat{A}_2) = E(z_0) = E(A + w_0) = A + 0 = A}$$

Mindkét becslő torzítatlan! $\boxed{\text{Melyik jobb?}}$

Számítsuk ki a becslők **varianciáját**!

$$var(\hat{A}_1) = var[\frac{1}{N}\sum_{k=0}^{N-1} z_k] = E\left\{\left(\left(\frac{1}{N}\sum_{k=0}^{N-1} z_k\right) - E\left\{\frac{1}{N}\sum_{k=0}^{N-1} z_k\right\}\right)^2\right\} = E\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1}(z_k - A)\right)^2\right\} = \frac{1}{N^2}\sum_{k=0}^{N-1} var(z_k) = \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N},$$

$$var(\hat{A}_2) = var(z_0) = \sigma^2 > var(\hat{A}_1)$$

Megjegyzés: **Ha több torzítatlan** becslőnk van ugyanarra a paraméterre vonatkozóan, és ezek **egymástól független** adathalmazokból származnak, például $\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{N-1}$, akkor jobb becslő nyerhető ezek átlagolásával:

$$\hat{a} = \frac{1}{N}\sum_{k=0}^{N-1} \hat{a}_k \Rightarrow E(\hat{a}) = a$$

Ha a becslőknek azonos a varianciája, akkor:

$$var(\hat{a}) = \frac{1}{N^2}\sum_{k=0}^{N-1} var(\hat{a}_k) = \frac{1}{N^2} N var(\hat{a}_k) = \frac{var(\hat{a}_k)}{N}.$$

N növelésével a variancia csökken, ha $N \to \infty, \hat{a} \to a$.

Ha a becslés torzított, akkor ez nem áll elő, függetlenül attól, hogy hány becslőt átlagolunk!

3.3.2. A minimális variancia kritériuma A becslés leglogikusabb minősítő kritériuma:

Az átlagos négyzetes hiba (Mean Square Error, MSE/*mse*), azaz a becslő és a mérendő különbsége négyzetének várható értéke:

$$mse(\hat{a}) = E[(\hat{a} - a)^2]$$

Sajnos ez a megközelítés **nem működőképes**, mert a becslő függ az **ismeretlen** *a* **paramétertől**.

Ami ehhez közel áll, az a becslő varianciája, amit bevezethetünk az *mse* kifejezésébe:

$mse(\hat{a}) = E\{[\hat{a} - E[\hat{a}] + E[\hat{a}] - a]^2\} = E\{[\hat{a} - E[\hat{a}] + b(a)]^2\}$. Itt $b(a) = E[\hat{a}] - a$ a becslő torzítását (bias) jelöli.

A torzítás (distortion) a mért jel alakjának vagy egyéb jellemzőjének nemkívánt változása. Additív zajt vagy egyéb külső jel hatását nem tekintjük torzításnak. Ha mért jel kifejezhető egy $y(t) = f(x(t))$ bemenet-kimenet relációval, és az $f(\quad)$ függvény $f^{-1}(\quad)$ inverze létezik, akkor akár a bemenetet, akár a kimenetet ezzel torzítva a torzítás megszüntethető/kompenzálható. Ha az inverz nem létezik, akkor a megszüntetés nem lehetséges.
Ha bemenet-kimenet reláció approximálható például rendszeridentifikációs módszerrel, akkor a torzítás egy közelítő eliminálása/kompenzálása lehetséges.

A zavarás (disturbance) a mért jel alakjának vagy egyéb jellemzőjének nem kívánt perturbációja. Tipikusan az additív zajt és egyéb nem identifikált vagy nem identifikálható külső jelet tekintünk zavarásnak. A zavarás nem szüntethető meg/nem kompenzálható csak csökkenthető.
(Identifikálható külső jel hatása persze kompenzálható, ilyen például az aktív zajcsökkentés.)

Tovább alakítva: $mse(\hat{a}) = E\{[\hat{a} - E(\hat{a})]^2\} + 2b(a)E[\hat{a} - E(\hat{a})] + b^2(a) = var(\hat{a}) + b^2(a)$.

Torzítatlan becslés esetén az *mse* minimalizálásával ekvivalens a becslő varianciájának minimalizálása!

3.3.3. Minimális varianciájú, torzítatlan becslők

Kézenfekvő törekvés olyan eljárások kidolgozása, amelyek nem torzítanak, azaz $b(a) = 0$, továbbá varianciájuk a lehető legkisebb.
Keressük a minimális varianciájú, torzítatlan becslőket! (**Minimum Variance Unbiased Estimator, MVU Estimator**)

Általános esetben nem mondható el, hogy MVU becslő létezik!

Lehet, hogy **nincsen** torzítatlan becslő.
Lehet, hogy **egyik** becslő **se minimális varianciájú**.

Mit tehetünk?

Érdekes módon van módszer arra, hogy meghatározzuk, hogy mekkora az elvileg elérhető legkisebb variancia!

Ez az ún. **Cramer-Rao** alsó korlát. (**Cramer-Rao Lower Bound: CRLB**)

1. Ezt kiszámítjuk és ellenőrizzük, hogy a becslőnk varianciája ennek megfelel-e?
2. Korlátozzuk magunkat **lineáris torzítatlan becslőkre**, mert ezek varianciája el tudja érni ezt az alsó korlátot.

Cramer-Rao alsó korlát (Cramer-Rao Lower Bound: CRLB, skalár paraméter esete)

Bármely torzítatlan becslő esetében a becslés varianciájára alsó korlátot ad: $var(\hat{a}) \geq CRLB(a)$ A CRLB a függvénye.

Megmondja mi az elérhető legkisebb variancia. Ha egy becslővel elérjük ezt a varianciát, akkor a becslőt **hatásos**nak nevezzük.

Maga a CRLB elvezethet az MVU becslőhöz. A CRLB-re vonatkozó állítás:

Tételezzük fel, hogy a mérési adatok valószínűség sűrűségfüggvényére teljesül minden a esetében: $E\left[\frac{\partial \ln f(z;a)}{\partial a}\right] = 0$

Ez az ún. **regularitási feltétel**, amihez az integrálás és a deriválás felcserélhetősége kötődik.

$$E\left[\frac{\partial \ln f(z;a)}{\partial a}\right] = \int_{-\infty}^{\infty} \frac{\partial \ln f(z;a)}{\partial a} f(z;a)dz = \frac{\partial}{\partial a}\int_{-\infty}^{\infty} f(z;a)dz = \frac{\partial}{\partial a}1 = 0$$

Ekkor:

$$var(\hat{a}) \geq \left[-E\left(\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right)\right]^{-1}$$

$$E\left[\frac{\partial lnf(z;a)}{\partial a}\right]=0,\quad \boxed{var(\hat{a})\geq\left[-E\left(\frac{\partial^2 lnf(z;a)}{\partial a^2}\right)\right]^{-1}}$$

Egy nagyon érdekes állítás:

Egy olyan **torzítatlan becslő**, amelyik eléri a Cramer-Rao alsó korlátot **akkor és csak akkor** található, ha:

$$\boxed{\frac{\partial lnf(z;a)}{\partial a}=I(a)(g(z)-a)}$$

struktúrájú, ahol $g(z)$ és $I(a)$ alkalmas függvények. Ilyenkor a becslő $\boxed{\hat{a}=g(z)}$ a variancia minimum: $\frac{1}{I(a)}$.

1. Példa: Additív, Gauss eloszlású fehér zajjal terhelt DC szint mérése egyetlen mért adatra alapozva: $\boxed{z_0=A+w_0,}$ ahol a w_0 valószínűség sűrűségfüggvénye $\mathcal{N}(0,\sigma^2)$.

$$f(z_0;A)=\frac{1}{\sqrt{2\pi\sigma^2}}exp\left[-\frac{1}{2\sigma^2}(z_0-A)^2\right],$$
$$lnf(z_0;A)=-ln\sqrt{2\pi\sigma^2}-\frac{1}{2\sigma^2}(z_0-A)^2,$$
$$\boxed{\frac{\partial\ln f(z_0;A)}{\partial A}=\frac{1}{\sigma^2}[z_0-A]}\Rightarrow\boxed{-\frac{\partial^2 lnf(z_0;A)}{\partial A^2}=\frac{1}{\sigma^2}.}$$

$$\boxed{\hat{A}=g(z_0)=z_0}$$
$$I(A)=\frac{1}{\sigma^2},\ \boxed{var(\hat{A})\geq\sigma^2}$$

2. Példa: Additív, Gauss eloszlású fehér zajjal terhelt DC szint mérése mérési sorozatra alapozva: $\boxed{z_k=A+w_k,}$ $k=0,1,\dots,N-1$

ahol a w_k korrelálatlan minden mintával, és valószínűségsűrűség függvénye $\mathcal{N}(0,\sigma^2)$.

$$\boxed{\hat{A}=\frac{1}{N}\sum_{k=0}^{N-1}z_k,}$$

$$f(z;A)=\frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^N}exp\left[-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(z_k-A)^2\right],$$
$$lnf(z;A)=-ln\left[(2\pi\sigma^2)^{\frac{N}{2}}\right]-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(z_k-A)^2,$$
$$\boxed{\frac{\partial lnf(z;A)}{\partial A}=\frac{1}{\sigma^2}\sum_{k=0}^{N-1}(z_k-A)}\Rightarrow\boxed{-\frac{\partial^2 lnf(z;A)}{\partial A^2}=\frac{N}{\sigma^2}.}\quad\boxed{var(\hat{A})\geq\frac{\sigma^2}{N}}$$

Az MVU becslő a mért értékek **átlaga**, a becslés varianciája pedig a mérőcsatornában fellépő additív zaj varianciájának ***N*-ed** része.

3. Példa: Ismeretlen variancia varianciájának alsó korlátja: $z_k=A+w_k, k=0,1,\dots,N-1;$ $\{w_k\}$ Gauss eloszlású, fehér zaj, ismeretlen σ^2 varianciával.

$$lnf(z;\sigma^2)=-\frac{N}{2}ln2\pi-\frac{N}{2}ln\sigma^2-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(z_k-A)^2,\quad\frac{\partial lnf(z;\sigma^2)}{\partial\sigma^2}=-\frac{N}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{k=0}^{N-1}(z_k-A)^2,$$

$$I(\sigma^2)=-E\left[\frac{\partial^2 lnf(z;\sigma^2)}{\partial(\sigma^2)^2}\right]=-E\left[\frac{N}{2\sigma^4}-\frac{\sum_{k=0}^{N-1}(z_k-A)^2}{\sigma^6}\right]=-\frac{N}{2\sigma^4}+\frac{N\sigma^2}{\sigma^6}=\frac{N}{2\sigma^4},$$

Innen a variancia alsó korlát: $\boxed{var(\widehat{\sigma^2})\geq\frac{2\sigma^4}{N}}$

Elérhetjük-e a CRLB értéket?

$$\boxed{\frac{\partial \ln f(z;a)}{\partial a} = I(a)(g(z)-a)}$$ **?** $$\frac{\partial \ln f(z;\sigma^2)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{k=0}^{N-1}(z_k - A)^2 = \frac{N}{2\sigma^4}\Big[\underbrace{\frac{1}{N}\sum_{k=0}^{N-1}(z_k - A)^2}_{\widehat{\sigma^2}} - \sigma^2\Big]$$

Igen, elérhetjük!

4. Példa: Ismeretlen a jelparaméter varianciájának alsó korlátja.

A megfigyelt értékek egy a-tól függő jel zajos mintái: $x_k = s(k;a) + w_k = s_k(a) + w_k$, $k = 0,1,\dots,N-1$; $\{w_k\}$ Gauss eloszlású, fehér zaj ismert σ_w^2 varianciával.

$$f(z;a) = \frac{1}{(2\pi\sigma_w^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_w^2}\sum_{k=0}^{N-1}(x_k - s_k(a))^2}, \quad \frac{\partial \ln f(z;a)}{\partial a} = \frac{1}{\sigma_w^2}\sum_{k=0}^{N-1}(x_k - s_k(a))\frac{\partial s_k(a)}{\partial a},$$

$$\frac{\partial^2 \ln f(z;a)}{\partial a^2} = \frac{1}{\sigma_w^2}\sum_{k=0}^{N-1}\left[(x_k - s_k(a))\frac{\partial^2 s_k(a)}{\partial a^2} - \left[\frac{\partial s_k(a)}{\partial a}\right]^2\right],$$ képezve a várható értéket:

$$I(a) = -E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right] = \frac{1}{\sigma_w^2}\sum_{k=0}^{N-1}\left[\frac{\partial s_k(a)}{\partial a}\right]^2, \quad \boxed{var(\hat{a}) \geq \frac{\sigma_w^2}{\sum_{k=0}^{N-1}\left[\frac{\partial s_k(a)}{\partial a}\right]^2}.}$$

Példa $s_k(a)$-ra: $A\cos(2\pi k\Delta t + \varphi)$
Paraméter lehet: $A, \varphi, \Delta t$

5. Példa: Függvénykapcsolattal származtatott $\alpha = g(a)$ paraméter varianciájának alsó korlátja: Bizonyítás a jegyzetben.

$$\boxed{var(\hat{\alpha}) \geq \frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left\{\left[\frac{\partial \ln f(z;a)}{\partial a}\right]^2\right\}} = -\frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right]}.}$$

Az $E\left\{\left[\frac{\partial \ln f(z;a)}{\partial a}\right]^2\right\} = -E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right]$ összefüggés a regularitási feltétel felhasználásával bizonyítható. Bizonyítás a jegyzetben.

6. Példa: Additív, Gauss eloszlású fehér zajjal terhelt DC szint (A) teljesítményének (A^2) „mérése" a DC szint mérésére alapozva.

A teljesítménymérés varianciájának alsó korlátja N minta figyelembevételével (a 2. példa eredményének felhasználásával):

Teljesítmény= A^2
Teljesítmény deriváltja= $2A$

$$\boxed{var(\hat{\alpha}) = var(\widehat{A^2}) \geq \frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left\{\left[\frac{\partial \ln f(z;a)}{\partial a}\right]^2\right\}} = -\frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right]} = \frac{(2A)^2}{\frac{N}{\sigma^2}} = \frac{4A^2\sigma^2}{N}}$$

Kérdés: Míg a DC szint hatásos becslője a minták egyszerű átlaga, vajon A^2 hatásos becslője-e $\hat{A}^2$? A válasz nemleges.

Ráadásul egy ilyen becslő **nem is torzítatlan,** ugyanis mivel $\hat{A} \sim \mathcal{N}(A, \sigma^2/N)$, a $var(\hat{A})$ kifejezéséből kiindulva:

$$E\{\hat{A}^2\} = E^2(\hat{A}) + var(\hat{A}) = A^2 + \frac{\sigma^2}{N} \neq A^2.$$

Egy Gauss eloszlású valószínűségi változó $\mathcal{N}(\mu, \sigma^2)$ momentumai explicit formában megadhatók: $E(a) = \mu,\ E(a^2) = \mu^2 + \sigma^2,\ E(a^3) = \mu^3 + 3\mu\sigma^2,\ E(a^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4, \ldots$

Vizsgáljuk meg $\hat{A}^2$ varianciáját! Mivel $var(\hat{A}^2) = E(\hat{A}^4) - E^2(\hat{A}^2)$, ahol

$$E(\hat{A}^4) = A^4 + 6A^2\frac{\sigma^2}{N} + 3\left(\frac{\sigma^2}{N}\right)^2,$$

$$var(\hat{A}^2) = A^4 + 6A^2\frac{\sigma^2}{N} + 3\left(\frac{\sigma^2}{N}\right)^2 - \left(A^2 + \frac{\sigma^2}{N}\right)^2 = \frac{4A^2\sigma^2}{N} + \frac{2\sigma^4}{N^2} > CRLB$$

A javasolt becslő nem lesz hatásos.

Cramer-Rao alsó korlát (Cramer-Rao Lower Bound: CRLB, vektor-paraméter esete)

Tegyük fel, hogy a mérési (vektor)adatok valószínűség sűrűségfüggvényére teljesül a regularitási feltétel minden $\boldsymbol{a}$ esetében:

$$E\left[\frac{\partial \ln f(z;\boldsymbol{a})}{\partial \boldsymbol{a}}\right] = 0.$$

Ekkor bármely torzítatlan becslőre igaz, hogy a becslő kovariancia mátrixának és az ún. **Fisher információs mátrix** inverzének a különbsége pozitív szemidefinit: $\boldsymbol{C}_{\hat{\boldsymbol{a}}} - \boldsymbol{I}^{-1}(\boldsymbol{a}) \geq 0$

A Fisher információs mátrix elemei:

$$\boldsymbol{I}_{ij}(\boldsymbol{a}) = -E\left(\frac{\partial^2 \ln f(z;\boldsymbol{a})}{\partial a_i \partial a_j}\right)$$

Egy olyan torzítatlan becslő, amelyik eléri a Cramer-Rao alsó korlátot akkor és csak akkor található, ha:

$$\frac{\partial \ln f(z;\boldsymbol{a})}{\partial \boldsymbol{a}} = \boldsymbol{I}(\boldsymbol{a})(\boldsymbol{g}(z) - \boldsymbol{a})$$

Ilyenkor a becslő $\hat{\boldsymbol{a}} = \boldsymbol{g}(z)$, és a kovariancia minimum $\boldsymbol{I}^{-1}(\boldsymbol{a})$.

7. Példa: Additív, Gauss eloszlású fehér zajjal terhelt DC szint és szórásának mérése mérési sorozatra alapozva:

$z_k = A + w_k$, $k = 0,1,\ldots,N-1$, ahol a w_k korrelálatlan minden mintával, és valószínűségsűrűség függvénye: $\mathcal{N}(0, \sigma^2)$.

Most $\boldsymbol{a} = [A \quad \sigma^2]^T$. A sűrűség függvény logaritmusa:

$$\ln f(z;\boldsymbol{a}) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(z_k - A)^2.$$

$$\boldsymbol{I}(\boldsymbol{a}) = -E\begin{bmatrix} \dfrac{\partial^2 \ln f(z;\boldsymbol{a})}{\partial A^2} & \dfrac{\partial^2 \ln f(z;\boldsymbol{a})}{\partial A \partial \sigma^2} \\ \dfrac{\partial^2 \ln f(z;\boldsymbol{a})}{\partial A \partial \sigma^2} & \dfrac{\partial^2 \ln f(z;\boldsymbol{a})}{\partial (\sigma^2)^2} \end{bmatrix} = \begin{bmatrix} \dfrac{N}{\sigma^2} & 0 \\ 0 & \dfrac{N}{2\sigma^4} \end{bmatrix}.$$

$$I(\boldsymbol{a}) = -E\begin{bmatrix} \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A^2} & \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A \partial \sigma^2} \\ \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A \partial \sigma^2} & \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial (\sigma^2)^2} \end{bmatrix} = \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}.$$

Ez a mátrix diagonális, így könnyen invertálható, amiből:

$$\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \begin{bmatrix} var(\hat{A}) & 0 \\ 0 & var(\widehat{\sigma^2}) \end{bmatrix} \geq \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix}$$

8. Példa:

Ha $\alpha = g(\boldsymbol{a})$ a becsülendő mennyiség, és az $\hat{\boldsymbol{a}}$ kovarianciája $\boldsymbol{I}^{-1}(\boldsymbol{a})$, akkor

$$\boldsymbol{C}_{\hat{\alpha}} \geq \left(\frac{\partial g(\boldsymbol{a})}{\partial \boldsymbol{a}}\right)\left[-E\left[\frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial \boldsymbol{a}^2}\right]\right]^{-1}\left(\frac{\partial g(\boldsymbol{a})}{\partial \boldsymbol{a}}\right)^T,$$

ahol a függvénykapcsolat paraméter szerinti deriváltját sorvektorként értelmezzük.

Ha az előző példához kapcsolódóan az $\alpha = g(\boldsymbol{a}) = A^2/\sigma^2$ jel/zaj viszony becslése varianciájának alsó határát keressük, akkor

$$\frac{\partial g(\boldsymbol{a})}{\partial \boldsymbol{a}} = \begin{bmatrix} \frac{\partial g(\boldsymbol{a})}{\partial A} & \frac{\partial g(\boldsymbol{a})}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix},$$

Ezt felhasználva a keresett alsó határ:

$$var(\hat{\alpha}) \geq \begin{bmatrix} \frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4} \end{bmatrix}\begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix}\begin{bmatrix} \frac{2A}{\sigma^2} \\ -\frac{A^2}{\sigma^4} \end{bmatrix} = \frac{4A^2}{N\sigma^2} + \frac{2A^4}{N\sigma^4} = \frac{4\alpha + 2\alpha^2}{N}$$

Megjegyzés:

Az előzőekben bemutatott példák közös jellemzője, hogy **determinisztikus paraméter**, és **ismert** valószínűségsűrűség függvényű **csatorna-karakterisztika** feltételezésével a becslés varianciájának **elvi minimumát** kapjuk meg.

Az, hogy ezt **milyen mérési eljárás** alkalmazása esetén érjük el, ill. **létezik-e, ismert-e** ilyen eljárás, sok esetben **nyitott kérdés.**

A továbbiakban korlátozzuk magunkat arra az esetre, amikor a **megfigyelési modell lineáris**.

Látni fogjuk, hogy ilyenkor hatásos becslőkhöz jutunk, azaz minden esetben elérjük a becslés varianciájára meghatározható alsó korlátot (CRLB).