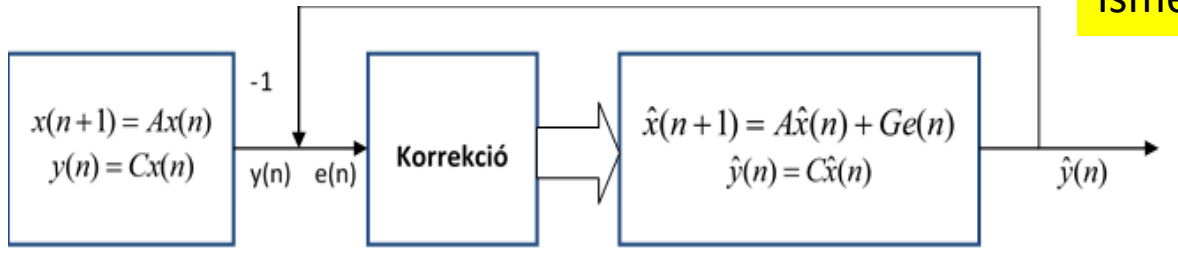
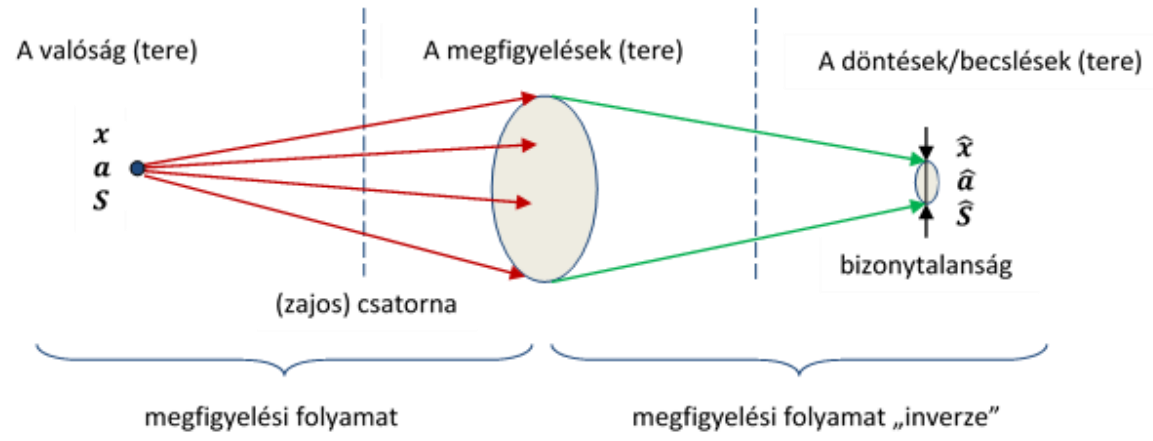




# Méréselmélet

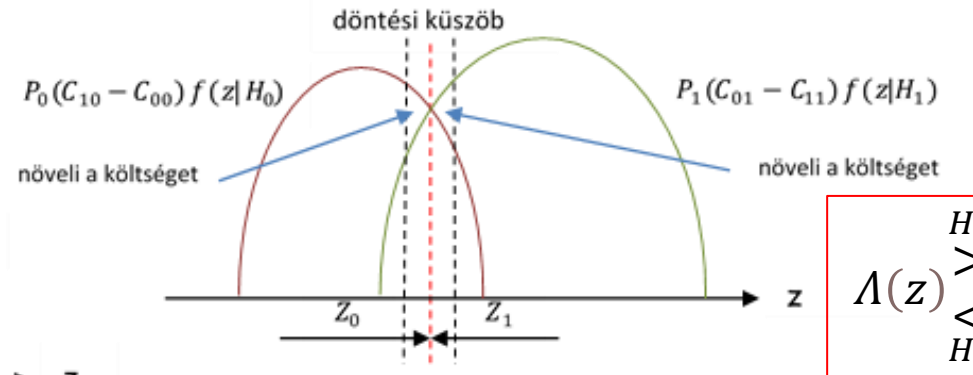
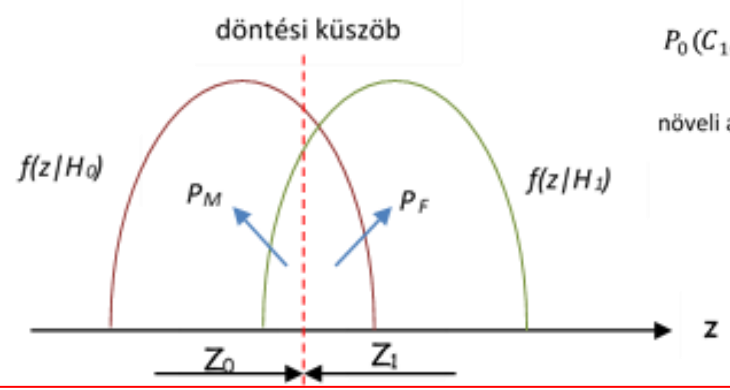
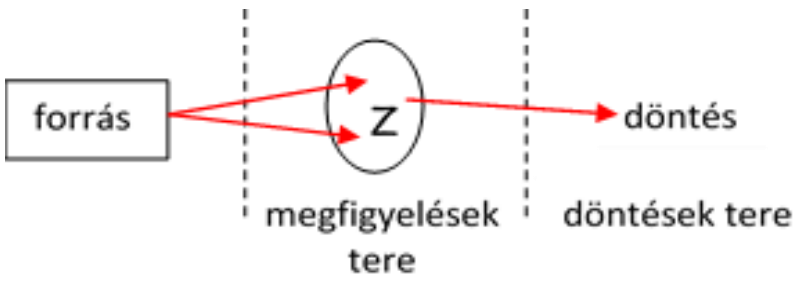
3. fejezet: A becslélmélet alapjai

2021. február 24.



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

## 2. A döntéelmélet alapjai



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

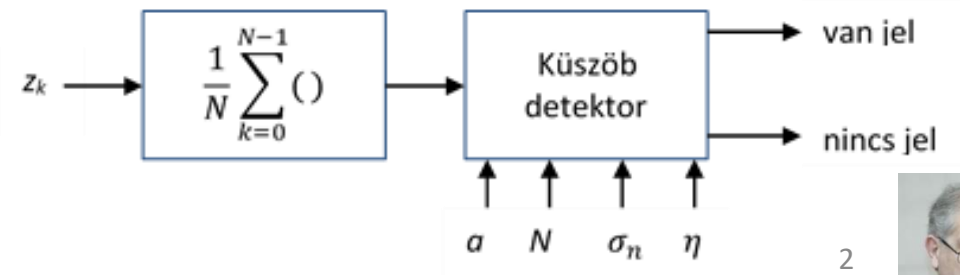
$$R = C_{10}P_0 + C_{11}P_1 + P_0(C_{00} - C_{10}) \int_{z_0} f(z|H_0)dz + P_1(C_{01} - C_{11}) \int_{z_0} f(z|H_1)dz.$$

$$\Lambda(z) = \frac{f(z|H_1)}{f(z|H_0)} \Big|_{z_{\text{küszöb}}} = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} = \eta$$

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

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

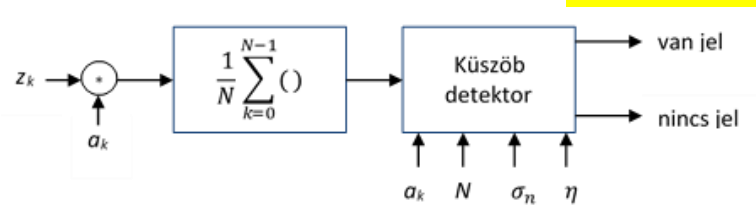
$$\frac{1}{N} \sum_{k=0}^{N-1} z_k > \frac{\sigma_n^2}{Na} \ln \eta + \frac{a}{2},$$



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

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

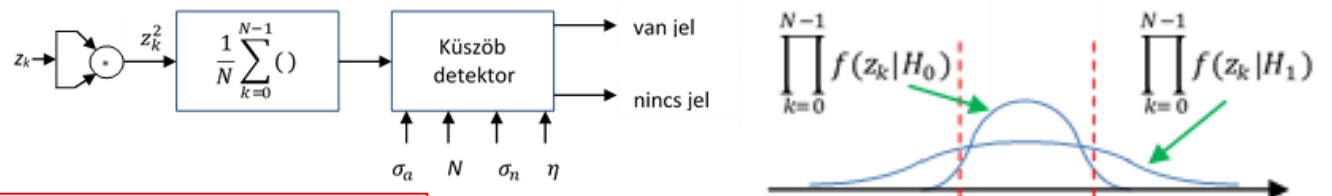
$$\frac{1}{N} \sum_{k=0}^{N-1} a_k z_k > \frac{\sigma_n^2}{N} \ln \eta + \frac{1}{2N} \sum_{k=0}^{N-1} a_k^2, \\ < \frac{\sigma_n^2}{N} \ln \eta + \frac{1}{2N} \sum_{k=0}^{N-1} a_k^2, H_0$$



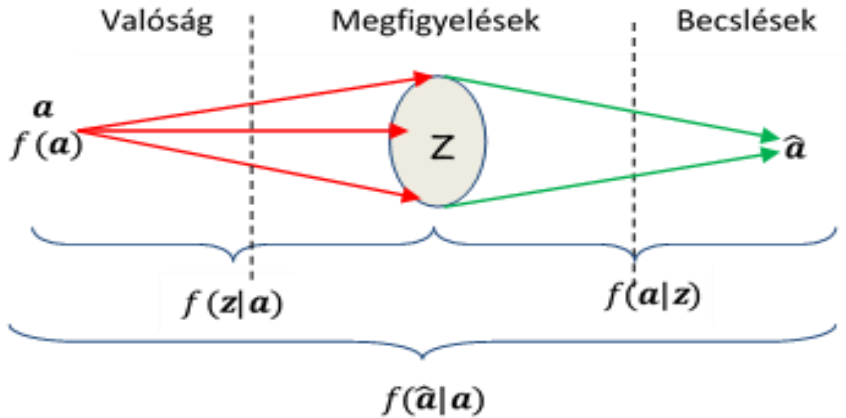
3. példa: Véletlen amplitúdójú jel detektálása: jelen van-e a jel a megfigyelésekben vagy nincsen?

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

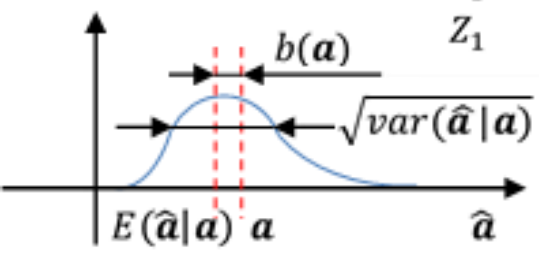
$$\frac{1}{N} \sum_{k=0}^{N-1} z_k^2 > \frac{2\sigma_n^2(\sigma_a^2 + \sigma_n^2)}{\sigma_a^2} \left[ \frac{1}{2} \ln \frac{\sigma_a^2 + \sigma_n^2}{\sigma_n^2} + \frac{1}{N} \ln \eta \right] \\ < \frac{2\sigma_n^2(\sigma_a^2 + \sigma_n^2)}{\sigma_a^2} \left[ \frac{1}{2} \ln \frac{\sigma_a^2 + \sigma_n^2}{\sigma_n^2} + \frac{1}{N} \ln \eta \right], H_0$$



3. A becslésemélet alapjai

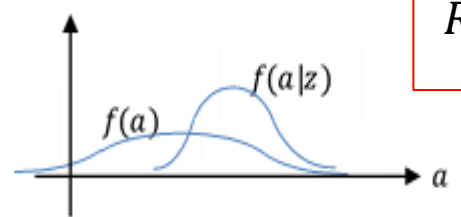


A cél az **a** paraméter(vektor) **â** becslőjének meghatározása az ábrán látható előzetesen ismert (és az azokból származtatott) sűrűségfüggvények ismeretében.



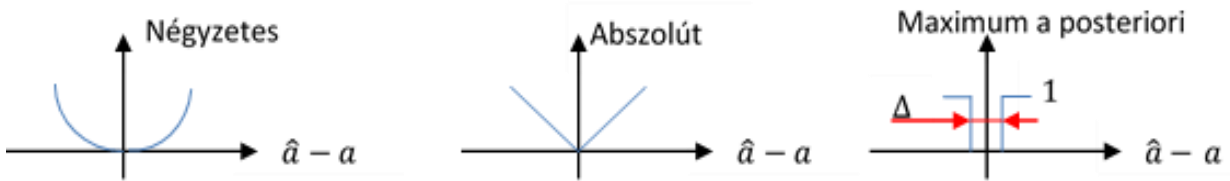
3.1. Bayes becslők

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



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

Itt **C(â, a)** az ún. költségfüggvény, vagy kockázatfüggvény, ...



3.1.1. Minimális átlagos négyzetes hibájú becslés

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

A legjobb becslés az  $a$  posteriori várható érték!

3.1.2. Minimális átlagos abszolút hibájú becslés

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

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

3.1.3. Maximum a posteriori (MAP) becslés

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

3.1.4. Bayes becselő 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:

$$\mathbf{z} = \mathbf{U}\mathbf{a} + \mathbf{n}$$

A megfigyelés modellje lineáris!

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

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

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

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

$$\text{cov}\{a, a|\mathbf{z}\} = \Sigma_{aa|z} = [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U} + \Sigma_{aa}^{-1}]^{-1}$$

**Megjegyzés:**  $\hat{a}_{MS} = \hat{a}_{ABS} = \hat{a}_{MAP}$ , 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.

" $U = IR + \text{zaj}$ " a modellünk. A megfigyelt értékek:  $z_k = a + n_k, k = 0, 1, \dots, N - 1$ ,  $a$  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**  $\mu_a$  és  $\sigma_a^2 \rightarrow$  az ismeretlen paraméter jellemzése.

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

Vektoros alakban:  $\mathbf{z} = \mathbf{U}\mathbf{a} + \mathbf{n}$   $\mathbf{z}^T = [z_0, z_2, \dots, z_{N-1}]$ ,  $\mathbf{n}^T = [n_0, n_2, \dots, n_{N-1}]$ ,  $\mathbf{U} = [1 \quad 1 \quad \dots \quad 1]^T$ ,  $\Sigma_{nn} = \sigma_n^2 \mathbf{I}$



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

$$\text{cov}\{a, a|\mathbf{z}\} = \Sigma_{aa|\mathbf{z}} = [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U} + \Sigma_{aa}^{-1}]^{-1}$$

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

Ha  $\sigma_a \ll \sigma_n$ , akkor  $\hat{a}_{MS} \cong \mu_a$ ,  
 ha  $\sigma_a \gg \sigma_n$  vagy  $N \rightarrow \infty$ , 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:

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

Ha  $\sigma_a \ll \sigma_n$ , akkor  $\text{var}\{\tilde{a}\} \cong \sigma_a^2$ ,  $\tilde{a} = \hat{a} - a$   
 ha  $\sigma_a \gg \sigma_n$  vagy  $N \rightarrow \infty$ , akkor  $\text{var}\{\tilde{a}\} \cong \frac{1}{N} \sigma_n^2$ .

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.

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

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

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

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

$$f(a|\mathbf{z}) = \frac{f(\mathbf{z}|a)f(a)}{f(\mathbf{z})} \quad \left. \frac{\partial \ln f(\mathbf{z})}{\partial a} = 0 \right. \quad f(\mathbf{z}|a) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_n^2} \sum_{k=0}^{N-1} (z_k - a s_k)^2} \rightarrow \frac{\partial \ln f(\mathbf{z}|a)}{\partial a} = \frac{1}{\sigma_n^2} \sum_{k=0}^{N-1} s_k (z_k - a s_k)$$





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

$$\frac{\partial \ln f(a)}{\partial a} = -\frac{a - \mu_a}{\sigma_a^2} \quad \frac{\partial \ln f(z|a)}{\partial a} = \frac{1}{\sigma_n^2} \sum_{k=0}^{N-1} s_k (z_k - a s_k)$$

$$\frac{1}{\sigma_n^2} \sum_{k=0}^{N-1} s_k (z_k - a s_k) - \frac{a - \mu_a}{\sigma_a^2} \Big|_{a=\hat{a}_{MAP}} = 0 \rightarrow$$

$$\hat{a}_{MAP} = \frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2} \sum_{k=0}^{N-1} s_k z_k}{1 + \frac{\sigma_a^2}{\sigma_n^2} \sum_{k=0}^{N-1} s_k^2}$$

$$var\{\tilde{a}\} = \frac{\sigma_a^2}{1 + \frac{\sigma_a^2}{\sigma_n^2} \sum_{k=0}^{N-1} s_k^2}$$

Ha  $s_k=1, \forall k$ , akkor megkapjuk az előző példa eredményét!

Itt is azonosítható a döntéseméleti rész 2. feladatában említett „illesztett” szűrő. Természetesen  $\hat{a}_{MAP} = \hat{a}_{MS}$ .

**Megjegyzés:** A variancia kifejezésének bizonyítása:

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

Itt  $E\{(\mu_a - a)^2\} = \sigma_a^2$   
 $E\{(\mu_a - a)(z_k - a s_k)\} = 0$   
 mert  $E\{n_k\} = 0$ , és  
 $E\{a n_k\} = 0, \forall k$

A második tag négyzete:

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

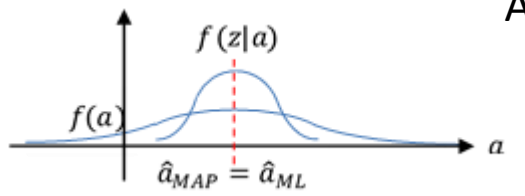
Ezzel:

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

### 3.2. Maximum Likelihood (ML) becslő

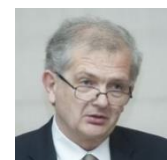
Kiindulásunk, hogy **nem ismerjük** a mérendő mennyiség a priori valószínűségi sűrűség függvényét. Ilyenkor azt feltételezzük, hogy ez a függvény „szélesen elterülő”, és ebből adódóan az *a posteriori* sűrűségfüggvény megegyezik a csatorna-karakterisztikával

Az optimális becslőt a csatorna-karakterisztika maximumához rendeljük:



$$\left. \frac{\partial f(z|a)}{\partial a} \right|_{a=\hat{a}_{ML}} = 0,$$

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



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

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

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

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

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

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

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

$$\mathbf{U}^T \Sigma_{nn}^{-1} (\mathbf{z} - \mathbf{U}\mathbf{a}) \Big|_{\mathbf{a}=\hat{\mathbf{a}}_{GM}} = 0 \rightarrow \hat{\mathbf{a}}_{GM} = [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1} \mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{z} \text{ ha } [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1} \text{ létezik.}$$

Várható értéke:  $\Sigma_{nn}$

**A becslés minőségjellemzője:**

$$cov(\hat{\mathbf{a}}_{GM}, \hat{\mathbf{a}}_{GM}) = E\{(\hat{\mathbf{a}}_{GM} - \mathbf{a})(\hat{\mathbf{a}}_{GM} - \mathbf{a})^T\} = E\{[\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1} \mathbf{U}^T \Sigma_{nn}^{-1} (\mathbf{z} - \mathbf{U}\mathbf{a})(\mathbf{z} - \mathbf{U}\mathbf{a})^T \Sigma_{nn}^{-1} \mathbf{U} [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1}\} = [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1}$$

$$cov(\hat{\mathbf{a}}_{GM}, \hat{\mathbf{a}}_{GM}) = [\mathbf{U}^T \Sigma_{nn}^{-1} \mathbf{U}]^{-1} \mathbf{I}$$

**Megjegyzések:**

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

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

$$\hat{\mathbf{a}}_{MS} = \hat{\mathbf{a}}_{GM}$$

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

**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:

$$f(\mathbf{z}|\mathbf{a}) = \frac{1}{(2\pi\sigma_n^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_n^2} \sum_{k=0}^{N-1} (z_k - a)^2}$$

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

$$\hat{a}_{ML} = \hat{a}_{GM} = \frac{1}{N} \sum_{k=0}^{N-1} z_k,$$

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

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

Ebben az esetben a „**csatorna-karakterisztika**”, azaz a  $z_0$  sűrűségfüggvénye:  $f(z_0; a) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(z_0 - a)^2\right]$  (Ez már nem feltételes,  $\mathbf{a}$  a függvény egy paramétere!)

**Példa:** Additív, **Gauss eloszlású fehér** zajjal terhelt **lineáris sorozat** mérése **mérési sorozatra** alapozva:  $z_k = A + Bk + w_k$ ,  $k = 0, 1, \dots, N - 1$ ,  $w_k$  korrelálatlan minden mintával, és valószínűség sűrűségfüggvénye  $\mathcal{N}(0, \sigma^2)$ .

Legyen  $\mathbf{a} = [A \ B]$ ,  $\mathbf{z} = [z_0, z_1, \dots, z_{N-1}]$ . Ebben az esetben a „**csatorna-karakterisztika**”, azaz  $\mathbf{z}$  sűrűségfüggvénye:

#### 3.3.1. Becslések minősítése

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

Tekintsünk egy  $A$  szintű DC mérést korrelálatlan zajban:

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

$$\hat{A}_1 = 0.95,$$

Tekintsük a következő becslőket:  $\hat{A}_1 = \frac{1}{N} \sum_{k=0}^{N-1} z_k$ ,  $\hat{A}_2 = z_0$ .

Tegyük fel, hogy  $A = 1$ ,

$$\hat{A}_2 = 0.98.$$

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:  $E(a - \hat{a}) = 0$

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

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

Mindkét becslő torzítatlan!

Melyik jobb?





Számítsuk ki a becslők **varianciáját!**

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

$\text{var}(\hat{A}_2) = \text{var}(z_0) = \sigma^2 > \text{var}(\hat{A}_1)$  **Megjegyzés:** Ha több torzítatlan becslőnk van ugyanarra a paraméterre vonatkozóan, és ezek **egymástól független** adathalmazokból származnak, például  $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{N-1}$ , 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:

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

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

**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:

$$\text{mse}(\hat{a}) = E[(\hat{a} - a)^2] \quad \text{Sajnos ez a megközelítés nem működőképes, mert a becslő függ az ismeretlen } a \text{ paramétertől.}$$

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

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

**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  $f(\cdot)$  függvény  $f^{-1}(\cdot)$  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, skálár 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 **hatásosnak** nevezzük.

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

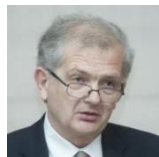
Tételezzük fel, hogy a mérési adatok valószínűség sűrűségfüggvényére teljesül minden  $a$  esetében:

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

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

Ekkor:

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



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

Egy nagyon érdekes állítás:

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

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

struktúrájú, ahol  $g(z)$  és  $I(a)$  alkalmas függvények.

Ilyenkor a becslő  $\hat{a} = g(z)$

$$\text{a variancia minimum: } \frac{1}{I(a)}.$$

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

$$z_0 = A + w_0,$$

$$\hat{A} = g(z_0) = z_0$$

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

$$\ln f(z_0; A) = -\ln \sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2} (z_0 - A)^2,$$

$$\frac{\partial \ln f(z_0; A)}{\partial A} = \frac{1}{\sigma^2} [z_0 - A]$$

$$\Rightarrow -\frac{\partial^2 \ln f(z_0; A)}{\partial A^2} = \frac{1}{\sigma^2}.$$

$$I(A) = \frac{1}{\sigma^2}, \quad \text{var}(\hat{A}) \geq \sigma^2$$

**2. Példa:** Additív, Gauss eloszlású fehér zajjal terhelt DC szint mérése mérési sorozatra alapozva: ahol a  $w_k$  korrelálatlan minden mintával, és valószínűségi sűrűség függvénye  $\mathcal{N}(0, \sigma^2)$ .

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

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

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

$$\ln f(z; A) = -\ln \left[ (2\pi\sigma^2)^{\frac{N}{2}} \right] - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} (z_k - A)^2,$$

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

$$\Rightarrow -\frac{\partial^2 \ln f(z; A)}{\partial A^2} = \frac{N}{\sigma^2}.$$

$$\text{var}(\hat{A}) \geq \frac{\sigma^2}{N}$$

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  $\sigma^2$  varianciával.

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

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

Innen a variancia alsó korlát:

$$\text{var}(\hat{\sigma}^2) \geq \frac{2\sigma^4}{N}$$



Elérhetjük-e a CRLB értéket?

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

?

$$\frac{\partial \ln f(z; \sigma^2)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{k=0}^{N-1} (z_k - A)^2 = \frac{N}{2\sigma^4} \left[ \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} (z_k - A)^2}_{\widehat{\sigma^2}} - \sigma^2 \right]$$

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  $\sigma_w^2$  varianciával.

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

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

$$I(a) = -E \left[ \frac{\partial^2 \ln f(z; a)}{\partial a^2} \right] = \frac{1}{\sigma_w^2} \sum_{k=0}^{N-1} \left[ \frac{\partial s_k(a)}{\partial a} \right]^2, \quad \text{Példa } s_k(a)\text{-ra: } A \cos(2\pi k \Delta t + \varphi)$$

$$\text{Paraméter lehet: } A, \varphi, \Delta t$$

$$\text{var}(\hat{a}) \geq \frac{\sigma_w^2}{\sum_{k=0}^{N-1} \left[ \frac{\partial s_k(a)}{\partial a} \right]^2}$$

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

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

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 =  $2A$

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



**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:  
 $E(a) = \mu, E(a^2) = \mu^2 + \sigma^2, E(a^3) = \mu^3 + 3\mu\sigma^2, E(a^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4, \dots$

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  $\mathbf{a}$  esetében:

$E \left[ \frac{\partial \ln f(z; \mathbf{a})}{\partial \mathbf{a}} \right] = 0.$  Ekkor bármely torzítatlan becslőre igaz, hogy a becslő kovariancia mátrixának és az ún. **Fisher információs mátrix** inverzének a különbsége pozitív szemidefinit:  $\mathbf{C}_{\hat{\mathbf{a}}} - \mathbf{I}^{-1}(\mathbf{a}) \geq 0$

A Fisher információs mátrix elemei:

$\mathbf{I}_{ij}(\mathbf{a}) = -E \left( \frac{\partial^2 \ln f(z; \mathbf{a})}{\partial a_i \partial a_j} \right)$  Egy olyan torzítatlan becslő, amelyik eléri a Cramer-Rao alsó korlátot akkor és csak akkor található, ha:  $\frac{\partial \ln f(z; \mathbf{a})}{\partial \mathbf{a}} = \mathbf{I}(\mathbf{a})(\mathbf{g}(z) - \mathbf{a})$

