Distribution of the GLRT for large data records and weak signal

Jessica Fridrich

EECE 566

Formulation

$$\begin{aligned} H_0: &\quad \mathbf{x} \sim p(\mathbf{x}; \boldsymbol{\theta}_{r_0}, \boldsymbol{\theta}_{s_0}) \\ H_1: &\quad \mathbf{x} \sim p(\mathbf{x}; \boldsymbol{\theta}_{r_1}, \boldsymbol{\theta}_{s_1}) \end{aligned}$$

- $\boldsymbol{\theta}_{r_i}$ parameter of **interest** under H_i
- $\boldsymbol{\theta}_{s_i}$ **nuisance** parameter under H_i

$$\text{(GLRT): } L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_{r_1}, \hat{\boldsymbol{\theta}}_{s_1}, H_1)}{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_{r_0}, \hat{\boldsymbol{\theta}}_{s_0}, H_0)} \underset{1}{\overset{0}{\lessgtr}} \gamma$$

- $\hat{\boldsymbol{\theta}}_{r_i}, \hat{\boldsymbol{\theta}}_{s_i}$ ML estimates of the parameters under H_i

Theorem (asymptotic distribution of GLRT)

1. $\boldsymbol{\theta}_{s_0} = \boldsymbol{\theta}_{s_1}$ nuisance parameters are the same under both hypotheses
2. $\|\boldsymbol{\theta}_{r_0} - \boldsymbol{\theta}_{r_1}\|$ small (Weak Signal)
3. n large (Large Data Sample, LDR)

$$2 \ln L_G(\mathbf{x}) \overset{(a)}{\sim} \begin{cases} \chi_r^2 & \text{under } H_0 \\ \chi_r'^2(\lambda) & \text{under } H_1 \end{cases}$$

λ non-centrality parameter (equivalent of deflection coefficient)
r dimensionality of $\boldsymbol{\theta}_{r_0}$ (the number of parameters of interest)
proof in App. 6A-C

Non-centrality parameter λ

Fisher information matrix

$$\begin{aligned}
\left[I(\boldsymbol{\theta})\right]_{i,j} &= -E_0\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln p(\mathbf{x};\boldsymbol{\theta})\right] \\
I(\boldsymbol{\theta}_r,\boldsymbol{\theta}_s) &= \begin{bmatrix} I_{\theta_r\theta_r}(\boldsymbol{\theta}_r,\boldsymbol{\theta}_r) & I_{\theta_r\theta_s}(\boldsymbol{\theta}_r,\boldsymbol{\theta}_s) \\ I_{\theta_s\theta_r}(\boldsymbol{\theta}_r,\boldsymbol{\theta}_s) & I_{\theta_s\theta_s}(\boldsymbol{\theta}_r,\boldsymbol{\theta}_s) \end{bmatrix}
\end{aligned}$$

$$\lambda = (\boldsymbol{\theta}_{r_0} - \boldsymbol{\theta}_{r_1})^T I(\lambda)(\boldsymbol{\theta}_{r_0} - \boldsymbol{\theta}_{r_1})$$

$I(\lambda) = I_{\theta_r\theta_r} - I_{\theta_r\theta_s} I^{-1}_{\theta_s\theta_s} I_{\theta_s\theta_r}$ is an $r \times r$ matrix evaluated at the true values $\boldsymbol{\theta}_{r_0}, \boldsymbol{\theta}_{s_0}$ under H_0

Corrolaries

- Pdf under H_0 does not depend on unknown parameters $\implies$ threshold can be determined for a given P_{FA}: Constant False-Alarm detector (CFAR)
- $I_{\theta_r\theta_r} \geq 0$, $I_{\theta_r\theta_s} I_{\theta_s\theta_s}^{-1} I_{\theta_s\theta_r} \geq 0$, both positive semi-definite (PSD) because they are parts of FI matrix, which is PSD $\implies$ the presence of $\boldsymbol{\theta}_s$ decreases λ because $\mathbf{x}^T I_{\theta_r\theta_s} I_{\theta_s\theta_s}^{-1} I_{\theta_s\theta_r} \mathbf{x} \geq 0$
  - recall $\lambda = \mathbf{x}^T I(\lambda)\mathbf{x}$, where $\mathbf{x} = \boldsymbol{\theta}_{r_0} - \boldsymbol{\theta}_{r_1}$
  - $I(\lambda) = I_{\theta_r\theta_r} - I_{\theta_r\theta_s} I_{\theta_s\theta_s}^{-1} I_{\theta_s\theta_r} \implies \lambda = \mathbf{x}^T I(\lambda)\mathbf{x} \leq \mathbf{x}^T I_{\theta_r\theta_r}\mathbf{x}$
- If no nuisance parameters, $\boldsymbol{\theta}_s = \{\emptyset\} \implies \boldsymbol{\theta} \equiv \boldsymbol{\theta}_r$ and λ increases, better detection

Example 1: Detection of unknown DC level in WGN (GLRT)

A **unknown**, σ^2 **known**, $\theta_{r_1} = \{A\}, \theta_{r_0} = \{0\}$, $\theta_{s_1} = \theta_{s_0} = \{\emptyset\}$
From previous lecture on GLRT:

$$2\ln L_G(\mathbf{x}) = \frac{n\overline{\mathbf{x}}^2}{\sigma^2}$$

$$\overline{\mathbf{x}} \sim \begin{cases} \mathcal{N}(0, \sigma^2/n) & \text{under } H_0 \\ \mathcal{N}(A, \sigma^2/n) & \text{under } H_1 \end{cases}$$

or normalized

$$\frac{\overline{\mathbf{x}}}{\sigma/\sqrt{n}} \sim \begin{cases} \mathcal{N}(0, 1) & \text{under } H_0 \\ \mathcal{N}\left(\frac{A}{\sigma/\sqrt{n}}, 1\right) & \text{under } H_1 \end{cases}$$

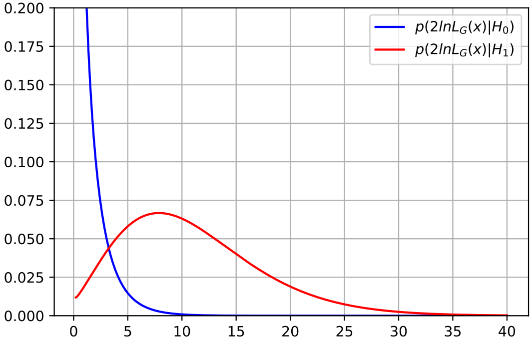
Example 1: Detection of unknown DC level in WGN (GLRT)

$$\left(\frac{\bar{\mathbf{x}}}{\sigma/\sqrt{n}}\right)^2 \sim \begin{cases} \chi_1^2 & \text{under } H_0 \\ \chi_1'^2(\lambda) & \text{under } H_1 \end{cases}$$

- Square of one Gaussian r.v. is χ_1^2 with 1 degree of freedom (pdf is derived in Problem 1 in assignment on statistics)
- The pdf's are **exact**, not just asymptotic in this special case
- Under H_1, $\chi_1'^2$ is the non-central chi-square with non-centrality $\lambda = \left(\frac{A}{\sigma/\sqrt{n}}\right)^2 = \frac{nA^2}{\sigma^2}$

Example 1: Detection of unknown DC level in WGN (GLRT)

$$n = 10,\ A = 1,\ \sigma^2 = 1$$



Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)

A **unknown**, σ^2 **unknown**, $\theta_{r_1} = \{A\}, \theta_{r_0} = \{0\}$,
$\theta_{s_1} = \theta_{s_0} = \{\sigma^2\}$
From previous lecture on GLRT:

$$2\ln L_G(\mathbf{x}) = n\ln\left(1 + \frac{\overline{\mathbf{x}}^2}{\hat{\sigma}_1^2}\right)$$

$$2\ln L_G(\mathbf{x}) \overset{(a)}{\sim} \begin{cases} \chi_1^2 & \text{under } H_0 \\ \chi_1'^2(\lambda) & \text{under } H_1 \end{cases}$$

- The distributions are only asymptotically chi-square
- The non-centrality parameter

$$\lambda = A^2\left[I_{AA}(0,\sigma^2) - I_{A\sigma^2}(0,\sigma^2)I^{-1}_{\sigma^2\sigma^2}(0,\sigma^2)I_{\sigma^2 A}(0,\sigma^2)\right]$$

Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)

- Gaussian FI matrix (see slides FI_Gaussian.pdf in Blackboard/Content)

$$I(A, \sigma^2) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}$$

- Non-centrality λ becomes (since $I_{A\sigma^2}(0, \sigma^2) = I_{\sigma^2 A}(0, \sigma^2) = 0$)

$$\begin{aligned} \lambda &= A^2 \left[I_{AA}(0, \sigma^2) - I_{A\sigma^2}(0, \sigma^2) I^{-1}_{\sigma^2\sigma^2}(0, \sigma^2) I_{\sigma^2 A}(0, \sigma^2) \right] \\ \lambda &= nA^2/\sigma^2 \end{aligned}$$

Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)

We can set the threshold (we did this before)

$$
\begin{aligned}
P_{\mathsf{FA}} &= \mathsf{Pr}\{2\ln L_G(\mathbf{x}) > \gamma | H_0\} \\
P_{\mathsf{FA}} &\overset{(a)}{=} \mathsf{Pr}\{\chi_1^2 > \gamma\} \\
&\overset{(a)}{=} \mathsf{Pr}\{\mathcal{N}(0,1)^2 > \gamma\} \\
&\overset{(a)}{=} \mathsf{Pr}\{|\mathcal{N}(0,1)| > \sqrt{\gamma}\} \\
&\overset{(a)}{=} \mathsf{Pr}\{\mathcal{N}(0,1) > \sqrt{\gamma}\} + \mathsf{Pr}\{\mathcal{N}(0,1) < -\sqrt{\gamma}\} \\
P_{\mathsf{FA}} &\overset{(a)}{=} 2\mathsf{Pr}\{\mathcal{N}(0,1) > \sqrt{\gamma}\} = 2Q(\sqrt{\gamma}) \\
\Rightarrow \gamma &\overset{(a)}{=} \left(Q^{-1}(P_{\mathsf{FA}}/2)\right)^2
\end{aligned}
$$

Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)

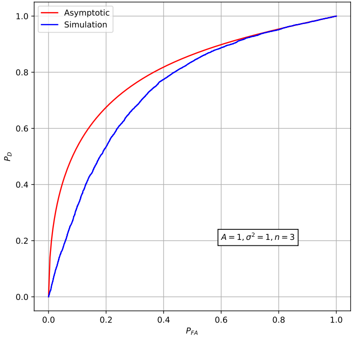
And determine the performance (did this before)

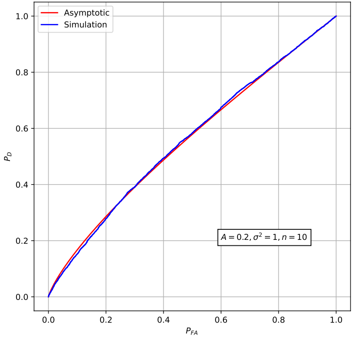
$$
\begin{aligned}
P_{\mathsf{D}} &= \mathsf{Pr}\{2\ln L_G(\mathbf{x}) > \gamma | H_1\} \\
P_{\mathsf{D}} &\overset{(a)}{=} \mathsf{Pr}\{\chi_1'^2 > \gamma\} \\
&\overset{(a)}{=} \mathsf{Pr}\{\mathcal{N}(\sqrt{\lambda}, 1)^2 > \gamma\} \\
&\overset{(a)}{=} \mathsf{Pr}\{|\mathcal{N}(\sqrt{\lambda}, 1)| > \sqrt{\gamma}\} \\
&\overset{(a)}{=} \mathsf{Pr}\{\mathcal{N}(\sqrt{\lambda}, 1) > \sqrt{\gamma}\} + \mathsf{Pr}\{\mathcal{N}(\sqrt{\lambda}, 1) < -\sqrt{\gamma}\} \\
&\overset{(a)}{=} Q(\sqrt{\gamma} - \sqrt{\lambda}) + Q(\sqrt{\gamma} + \sqrt{\lambda}) \\
P_{\mathsf{D}} &\overset{(a)}{=} Q\left(Q^{-1}(P_{\mathsf{FA}}/2) - \sqrt{\lambda}\right) + Q\left(Q^{-1}(P_{\mathsf{FA}}/2) + \sqrt{\lambda}\right)
\end{aligned}
$$

Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)

- How close is the **actual detector** performance to the **asymptotic approximation**?
  - Monte-Carlo sampling
  - Generate M samples under H_0 and M samples under H_1
  - Use $L_G(\mathbf{x})$ or $T(\mathbf{x}) = \frac{\overline{\mathbf{x}}^2}{\hat{\sigma}_1^2}$ (because $\ln(1+x)$ is monotone) and draw the ROC curve
  - Compare to the asymptotic ROC curve derived above $P_{\mathsf{D}} = Q\left(Q^{-1}(P_{\mathsf{FA}}/2) - \sqrt{\lambda}\right) + Q\left(Q^{-1}(P_{\mathsf{FA}}/2) + \sqrt{\lambda}\right)$

Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)





Example 2: Detection of unknown DC level in WGN of unknown variance (GLRT, LDR)