Optimum GLR Tests

George V. Moustakides
University of Patras, GREECE

Outline

- Hypothesis testing and GLR test
- Randomized tests
  - Classical binary hypothesis testing
  - Composite hypotheses and selection (isolation)
  - **Two different implementations of randomized tests**
- Combined hypothesis testing/isolation – Optimality of GLRT
- Randomized estimators
- Combined hypothesis testing/estimation – Optimum GLR tests

Hypothesis testing and GLRT

Binary hypothesis testing (fixed sample size N)

Collection of observations $X = [x_1, x_2, \ldots, x_N]$

$$\mathbb{H}_0 : \; X \; \sim \; f_0(X)$$
$$\mathbb{H}_1 : \; X \; \sim \; f_1(X)$$

Likelihood Ratio Test $$\frac{f_1(X)}{f_0(X)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda$$

Neyman-Pearson, Bayes, Min-max

Composite binary hypothesis testing

$X=[x_1,x_2,\ldots,x_N]$

$$\begin{aligned}\mathbb{H}_0 : X &\sim f_{01}(X)\\ &\vdots\\ &\sim f_{0K_0}(X)\\ \mathbb{H}_1 : X &\sim f_{11}(X)\\ &\vdots\\ &\sim f_{1K_1}(X)\end{aligned}$$

One subcase responsible for the data

If subcases were known say $f_{0i}(X)$ and $f_{1j}(X)$

$$\frac{f_{1j}(X)}{f_{0i}(X)} \mathop{\gtreqless}_{\mathbb{H}_0}^{\mathbb{H}_1} \lambda$$

In order to **isolate** the subcase responsible for X

$$\frac{\max_j\{f_{1j}(X)\}}{\max_i\{f_{0i}(X)\}} \mathop{\gtreqless}_{\mathbb{H}_0}^{\mathbb{H}_1} \lambda$$

GLRT

$$\frac{f_{1\hat{j}}(X)}{f_{0\hat{i}}(X)} \mathop{\gtreqless}_{\mathbb{H}_0}^{\mathbb{H}_1} \lambda \qquad \begin{aligned} \hat{i} &= \arg\max_i\{f_{0i}(X)\} \\ \hat{j} &= \arg\max_j\{f_{1j}(X)\} \end{aligned}$$

Detection with **Maximum Likelihood isolation**

$$\mathbb{H}_0 : \quad X \sim f_0(X|\theta_0)$$
$$\mathbb{H}_1 : \quad X \sim f_1(X|\theta_1)$$

Popular decision method in practice

$$\frac{\sup_{\theta_1}\{f_1(X|\theta_1)\}}{\sup_{\theta_0}\{f_0(X|\theta_0)\}} \mathop{\gtreqless}_{\mathbb{H}_0}^{\mathbb{H}_1} \lambda$$

GLRT

$$\frac{f_1(X|\hat{\theta}_1)}{f_0(X|\hat{\theta}_0)} \mathop{\gtreqless}_{\mathbb{H}_0}^{\mathbb{H}_1} \lambda$$

$$\hat{\theta}_0 = \arg\sup_{\theta_0}\{f_0(X|\theta_0)\}$$
$$\hat{\theta}_1 = \arg\sup_{\theta_1}\{f_1(X|\theta_1)\}$$

Detection with **Maximum Likelihood estimation**

Optimality of GLRT

Wald (1943)

Le Cam (1953)

Lehmann (1959)

Hoeffding (1965)

Neyman (1965)

⋮

Gabriel and Kay (2005)

Nicolls and De Jager (2007)

Asymptotic optimality under regularity conditions. Performance of GLRT approaches the optimum test (with true parameters known).

Randomized tests

Binary hypothesis testing (fixed sample size N)

Collection of observations $X=[x_1,x_2,\dots,x_N]$

$$\mathbb{H}_0 : X \sim f_0(X)$$

$$\mathbb{H}_1 : X \sim f_1(X)$$

Decide between $\mathbb{H}_0$ and $\mathbb{H}_1$

Deterministic tests

Two complementary sets A_0 , A_1 in $\mathbb{R}^N$

Decide $\mathbb{H}_i$ when $X \in A_i$

Randomized tests

Two complementary probabilities $\delta_0(X)$, $\delta_1(X)$

In a **random game** decide $\mathbb{H}_i$ with probability $\delta_i(X)$

Randomized is more general than deterministic

$$\delta_i(X) = 1_{A_i}(X)$$

Suppose $X \in A_0$ then with **probability 1** decide in favor of $\mathbb{H}_0$ (equivalent to deterministic decision)

Neyman-Pearson

$$\max_{\delta_0,\delta_1} \mathbb{P}[d = 1|\mathbb{H}_1] \qquad \mathbb{P}[d = 1|\mathbb{H}_0] \leq \alpha$$

$$\int \delta_1(X) f_1(X) dX - \lambda \int \delta_1(X) f_0(X) dX$$

$$\max_{\delta_0,\delta_1} \int \delta_1(X)[f_1(X) - \lambda f_0(X)] dX$$

$$\delta_1(X) = \begin{cases} 1 & f_1(X) - \lambda f_0(X) > 0 \\ \gamma & f_1(X) - \lambda f_0(X) = 0 \\ 0 & f_1(X) - \lambda f_0(X) < 0 \end{cases}$$

$$\delta_0(X) = \begin{cases} 0 & f_1(X) - \lambda f_0(X) > 0 \\ 1-\gamma & f_1(X) - \lambda f_0(X) = 0 \\ 1 & f_1(X) - \lambda f_0(X) < 0 \end{cases}$$

$$\frac{f_1(X)}{f_0(X)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda$$

Composite binary hypotheses

$$
\begin{array}{llll}
\mathbb{H}_0 : X & \sim & f_{01}(X),\ \pi_{01} & \delta_{01}(X) \\
 & & \vdots & \vdots \\
 & \sim & f_{0K_0}(X),\ \pi_{0K_0} & \delta_{0K_0}(X) \\
\mathbb{H}_1 : X & \sim & f_{11}(X),\ \pi_{11} & \delta_{11}(X) \\
 & & \vdots & \vdots \\
 & \sim & f_{1K_1}(X),\ \pi_{1K_1} & \delta_{1K_1}(X)
\end{array}
$$

$$\delta_{01}(X)\cdots\delta_{0K_0}(X) \qquad \delta_{11}(X)\cdots\delta_{1K_1}(X)$$

$$\sum_{i=0,1}\sum_{j=1}^{K_i}\delta_{ij}(X)=1$$

With a random game, in a **single step**, with probability $\delta_{ij}(X)$ decide in favor of $\mathbb{H}_i$ **and** isolate subcase j.

We **simultaneously** detect **and** isolate.

Alternative two step implementation

$$\underbrace{\delta_{01}(X)\cdots\delta_{0K_0}(X)}_{+} \qquad \underbrace{\delta_{11}(X)\cdots\delta_{1K_1}(X)}_{+}$$

$$\delta_0(X) \qquad \delta_1(X)$$

$$\delta_0(X)+\delta_1(X)=1$$

$$\gamma_{ij}(X)=\frac{\delta_{ij}(X)}{\delta_i(X)} \qquad \gamma_{i1}(X)+\cdots+\gamma_{iK_i}(X)=1$$

- With probabilities $\delta_0(X)$, $\delta_1(X)$ decide in favor of $\mathbb{H}_0$ or $\mathbb{H}_1$.
- Given we selected $\mathbb{H}_i$ use probabilities $\gamma_{i1}(X)$, $\gamma_{i2}(X)$,… to isolate one of the subcases of $\mathbb{H}_i$.

Optimality of GLRT

$$\mathbb{H}_0 : X \sim f_{01}(X), \pi_{01} \quad \vdots \quad \sim f_{0K_0}(X), \pi_{0K_0}$$

$$\mathbb{H}_1 : X \sim f_{11}(X), \pi_{11} \quad \vdots \quad \sim f_{1K_1}(X), \pi_{1K_1}$$

Combined Detection and Isolation

Following a Neyman-Pearson like approach:

$$\max \mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_1]$$

subject to the "false alarm" constraint

$$\mathbb{P}[\text{Miss-detection/isolation}|\mathbb{H}_0] \leq \alpha$$

Theorem: The test that solves the combined detection/isolation problem is

$$\frac{\max\limits_{1\le j\le K_1}\{\pi_{1j}f_{1j}(X)\}}{\max\limits_{1\le j\le K_0}\{\pi_{0j}f_{0j}(X)\}} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda$$

Threshold λ and randomization probability γ are selected to satisfy the constraint with equality.

If priors unknown, use equiprobable

$$\Rightarrow \frac{\max\limits_{1\le j\le K_1}\{f_{1j}(X)\}}{\max\limits_{1\le i\le K_0}\{f_{0i}(X)\}} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda\frac{K_1}{K_0}$$

Proof:

$$\mathbb{P}[\text{Miss-detection/isolation}|\mathbb{H}_0] \leq \alpha \quad \Leftrightarrow$$

$$\mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_0] \geq 1-\alpha$$

$$\mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_i]$$

$$= \sum_{j=1}^{K_i} \mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_{ij}]\pi_{ij}$$

$$\mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_{ij}]$$

$$= \int \delta_i(X)\gamma_{ij}(X)f_{ij}(X)dX$$

$$\mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_1] + \lambda\,\mathbb{P}[\text{Correct-detection/isolation}|\mathbb{H}_0]$$

$$= \int \delta_1(X)\left\{\sum_{j=1}^{K_1}\gamma_{1j}(X)\pi_{1j}f_{1j}(X)\right\}dX + \lambda\int\delta_0(X)\left\{\sum_{j=1}^{K_0}\gamma_{0j}(X)\pi_{0j}f_{0j}(X)\right\}dX$$

$$\leq \int \delta_1(X)\max_{1\leq j\leq K_1}\{\pi_{1j}f_{1j}(X)\}\,dX + \lambda\int\delta_0(X)\max_{1\leq j\leq K_0}\{\pi_{0j}f_{0j}(X)\}\,dX$$

$$= \int \Big[\delta_1(X) \max_{1\le j\le K_1} \{\pi_{1j} f_{1j}(X)\}$$

$$+ \delta_0(X)\lambda \max_{1\le j\le K_0} \{\pi_{0j} f_{0j}(X)\}\Big] dX$$

$$\le \int \max\Big\{ \max_{1\le j\le K_1} \{\pi_{1j} f_{1j}(X)\},$$

$$\lambda \max_{1\le j\le K_0} \{\pi_{0j} f_{0j}(X)\}\Big\} dX$$

$$\Leftrightarrow \frac{\max_{1\le j\le K_1} \{\pi_{1j} f_{1j}(X)\}}{\max_{1\le j\le K_0} \{\pi_{0j} f_{0j}(X)\}} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda$$

Randomized estimators

Collection of observations $X=[x_1,x_2,\dots,x_N]$

$$X\sim f(X|\theta),\ \pi(\theta)$$

We would like to obtain an estimate $\hat{\theta}$ of θ.

Deterministic estimator: $\hat{\theta}=G(X)$

Randomized estimator: Must assign probability to every possible value of $\hat{\theta}$.

$$\delta(\hat{\theta}|X)d\hat{\theta} \text{ (differential) } \int\delta(\hat{\theta}|X)d\hat{\theta}=1$$

A **pdf** with respect to $\hat{\theta}$ conditioned on X.

Generate r. v. $\hat{\theta}$ distributed according to $\delta(\hat{\theta}|X)$

"Deterministic" $\delta(\hat{\theta}|X)=\mathrm{Dirac}(\hat{\theta}-G(X))$

Optimum Bayes estimation

If true parameter θ and estimate $\hat{\theta}$ then cost $\mathcal{C}(\hat{\theta},\theta)$. **Minimize average cost.**

$$\mathbb{E}[\mathcal{C}(\hat{\theta},\theta)] = \iiint \mathcal{C}(\hat{\theta},\theta)\delta(\hat{\theta}|X)f(X|\theta)\pi(\theta)\,dX\,d\hat{\theta}\,d\theta$$

$$= \int\left[\int \delta(\hat{\theta}|X)\underbrace{\left\{\int \mathcal{C}(\hat{\theta},\theta)f(X|\theta)\pi(\theta)d\theta\right\}}_{\mathcal{D}(\hat{\theta},X)\,\geq\,\inf_U \mathcal{D}(U,X)} d\hat{\theta}\right]dX$$

$$\geq \int \inf_U \mathcal{D}(U,X)dX$$

$$\delta(\hat{\theta}|X) = \mathrm{Dirac}\left(\hat{\theta} - \arg\inf_U \mathcal{D}(U,X)\right)$$

Combined detection/estimation

Collection of observations $X=[x_1,x_2,...,x_N]$

$$\mathbb{H}_0 : \ X \sim f(X|\theta = 0)$$

$$\mathbb{H}_1 : \ X \sim f(X|\theta), \ \pi(\theta)$$

- Decide between $\mathbb{H}_0$ and $\mathbb{H}_1$
- Whenever I decide in favor of $\mathbb{H}_1$ must provide estimate $\hat{\theta}$ for θ.

For detection/estimation a two-step strategy:

$$\delta_0(X), \ \delta_1(X), \ \gamma(\hat{\theta}|X)$$

Following a Neyman-Pearson like approach we propose:

$$\inf \mathbb{E}[\mathcal{C}(\hat{\theta},\theta)|\mathbb{H}_1]$$

subject to: $\mathbb{P}[\text{Miss-detection}|\mathbb{H}_0] \leq \alpha$

$$\mathbb{P}[\text{Miss-detection}|\mathbb{H}_0] = \int \delta_1(X) f(X|0) dX$$

Average cost computation under $\mathbb{H}_1$

Wrong hypothesis:

$$\iint \mathcal{C}(0,\theta)\delta_0(X) f(X|\theta)\pi(\theta) d\theta dX$$

Correct hypothesis:

$$\iiint \mathcal{C}(\hat{\theta},\theta)\delta_1(X)\gamma(\hat{\theta}|X) f(X|\theta)\pi(\theta) d\hat{\theta} d\theta dX$$

$$\iiint \mathcal{C}(\hat\theta,\theta)\delta_1(X)\gamma(\hat\theta|X)f(X|\theta)\pi(\theta)d\hat\theta d\theta dX$$
$$+\iint \mathcal{C}(0,\theta)\delta_0(X)f(X|\theta)\pi(\theta)d\theta dX$$
$$+\lambda\int\delta_1(X)f(X|0)dX$$

$$\mathcal{D}(\hat\theta,X)\geq\inf_U\mathcal{D}(U,X)$$

$$=\int\delta_1(X)\Big[\int\gamma(\hat\theta|X)\Big\{\underbrace{\int\mathcal{C}(\hat\theta,\theta)f(X|\theta)\pi(\theta)d\theta}\Big\}d\hat\theta\Big]dX$$
$$+\int\delta_0(X)\Big\{\underbrace{\int\mathcal{C}(0,\theta)f(X|\theta)\pi(\theta)d\theta}_{\mathcal{D}(0,X)}\Big\}dX$$
$$+\lambda\int\delta_1(X)f(X|0)dX$$

$$\geq\int\delta_1(X)\inf_U\mathcal{D}(U,X)dX+\int\delta_0(X)\mathcal{D}(0,X)dX$$
$$+\lambda\int\delta_1(X)f(X|0)dX$$

$$\int \Big\{\delta_1(X)\Big[\inf_U \mathcal{D}(U,X) + \lambda f(X|0)\Big] + \delta_0(X)\mathcal{D}(0,X)\Big\}dX$$

$$\geq \int \min\{\mathcal{D}(U,X) + \lambda f(X|0), \mathcal{D}(0,X)\}dX$$

$$\frac{\mathcal{D}(0,X) - \inf_U \mathcal{D}(U,X)}{f(X|0)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}} \lambda$$

$$\mathcal{D}(U,X) = \int \mathcal{C}(U,\theta) f(X|\theta)\pi(\theta)d\theta$$

Examples

MAP estimator

$$\mathcal{C}(\hat{\theta},\theta)=\begin{cases}0 & \|\hat{\theta}-\theta\|\le\Delta\ll 1\\ 1 & \text{otherwise}\end{cases}$$

Optimum test: $$\frac{\sup_{\theta}\pi(\theta)f(X|\theta)}{f(X|0)}\ \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}}\ \lambda$$

With uniform prior: $$\frac{\sup_{\theta}f(X|\theta)}{f(X|0)}\ \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtrless}}\ \lambda$$ **GLRT**

MMSE estimator

$$\mathcal{C}(\hat{\theta}, \theta) = \|\hat{\theta} - \theta\|^2$$

Optimum test:

$$\|\hat{\theta}_{\text{MMSE}}\|^2 \frac{\int f(X|\theta)\pi(\theta)d\theta}{f(X|0)} \underset{\mathbb{H}_0}{\overset{\mathbb{H}_1}{\gtreqless}} \lambda$$

$$\hat{\theta}_{\text{MMSE}} = \mathbb{E}[\theta|X] = \frac{\int \theta\pi(\theta)f(X|\theta)d\theta}{\int \pi(\theta)f(X|\theta)d\theta}$$

Median estimator

$$\mathcal{C}(\hat\theta,\theta)=|\hat\theta-\theta|$$

Optimum test:

$$\frac{\int_0^{\hat\theta_{\rm MEDIAN}}\theta\pi(\theta)f(X|\theta)d\theta}{f(X|0)}\ \overset{\mathbb{H}_1}{\underset{\mathbb{H}_0}{\gtreqless}}\ \lambda$$

$$\hat\theta_{\rm MEDIAN}=\arg\left\{\hat\theta:\mathbb{P}[\theta\le\hat\theta|X]=0.5]\right\}$$

$$=\arg\left\{\hat\theta:\frac{\int_{-\infty}^{\hat\theta}\pi(\theta)f(X|\theta)d\theta}{\int_{-\infty}^{\infty}\pi(\theta)f(X|\theta)d\theta}=0.5\right\}$$