BAYES FACTORS AND GLRTS

Signal Detection in the presence of unknowns: Part II

We've seen that in special cases of parametric uncertainty (one-sided tests with monotonic likelihood ratios), the LRT reduces to a UMP thresholding test.

Generally, UMP test do not exist. We will study two popular methods for devising (usually) sub-optimal detectors in these more challenging problems:

1. Bayes factors
2. Generalized LRTs

These two methods differ in how they model the unknown parameter $\underline{\theta}$:

1. $\underline{\theta}$ is itself a random quantity
   (Bayesian approach)
2. $\underline{\theta}$ is unknown, but fixed
   (Classical approach)

Bayes Factors

Consider

$$H_0 : \underline{X} \sim f(\underline{x} \mid \underline{\theta}_0)$$
$$H_1 : \underline{X} \sim f(\underline{x} \mid \underline{\theta}_1)$$

not necessarily same parametric family

where $\underline{\theta}_0$ and $\underline{\theta}_1$ are unknown.

Assume $\underline{\theta}_0$ and $\underline{\theta}_1$ are realizations of the prior distributions

$$f(\underline{\theta}_k \mid H_k), \quad k = 0, 1$$

Then

$$f(\underline{x} \mid H_k) = \int f(\underline{x}, \underline{\theta}_k \mid H_k)\, d\underline{\theta}_k$$

$$= \int f(\underline{x} \mid H_k, \underline{\theta}_k)\, f(\underline{\theta}_k \mid H_k)\, d\underline{\theta}_k$$

Thus the LR statistic is

$$\Lambda(\underline{x}) = \frac{f(\underline{x} \mid H_1)}{f(\underline{x} \mid H_0)}$$

$$= \frac{\int f(\underline{x} \mid H_1, \underline{\theta}_1)\, f(\underline{\theta}_1 \mid H_1)\, d\underline{\theta}_1}{\int f(\underline{x} \mid H_0, \underline{\theta}_0)\, f(\underline{\theta}_0 \mid H_0)\, d\underline{\theta}_0}$$

This "integrated likelihood ratio" is called the <u>Bayes factor</u> for testing H_1 vs. H_0

Example: Geometric vs. Poisson

H_0 : $x_1, \dots, x_N$ iid geometric

$$x_i \sim \theta_0 (1-\theta_0)^{x_i} , \qquad 0 \le \theta_0 \le 1$$

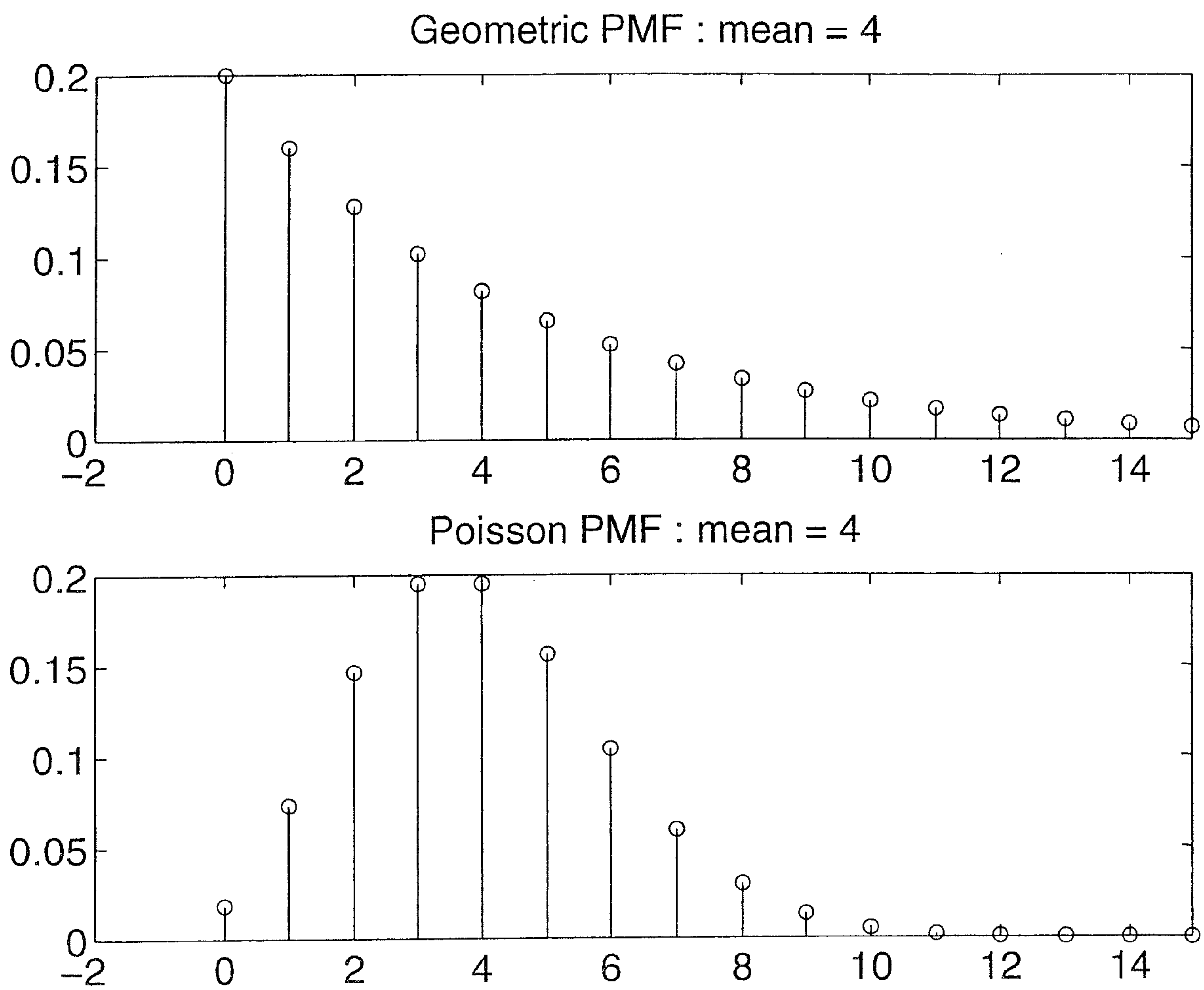
$$f(\underline{x} \mid H_0, \theta_0) = \theta_0^N (1-\theta_0)^{\sum_{i=1}^N x_i}$$

H_1 : $x_1, \dots, x_N$ iid Poisson

$$x_i \sim e^{-\theta_1} \frac{\theta_1^{x_i}}{x_i!} , \qquad \theta_1 \ge 0$$

$$f(\underline{x} \mid H_1, \theta_1) = e^{-N\theta_1} \frac{\theta_1^{\sum_{i=1}^N x_i}}{\prod_{i=1}^N x_i!}$$

Can we apply Karlin-Rubin?


Geometric PMF : mean = 4
0.2
0.15
0.1
0.05
0
−2
0
2
4
6
8
10
12
14
Poisson PMF : mean = 4
0.2
0.15
0.1
0.05
0
−2
0
2
4
6
8
10
12
14

Taking a Bayesian approach, let's assume prior distributions for θ_0 and θ_1:

Modeling uncertainty in θ_k

Under H_0 : $0 \le \theta_0 \le 1$

$$f(\theta_0) = I_{[0,1]}(\theta_0)$$

uniform : no preference for any value

Under H_1 : $\theta_1 \geq 0$

$$f(\theta_1) = e^{-\theta_1}$$

exponential : favors smaller values

Integrated Likelihoods

$$f(\underline{x} \mid H_0) = \int_0^1 \theta^N (1-\theta)^t \, d\theta = \frac{N!\,t!}{(N+t+1)!}$$

$$t = \sum_{i=1}^{N} x_i$$

$$f(\underline{x} \mid H_1) = \int_0^\infty \frac{e^{-(N+1)\theta} \, \theta^t}{\prod_{i=1}^{N} x_i!} = \frac{t!}{(N+1)^{t+1} \prod_{i=1}^{N} x_i!}$$

Bayes Factor:

$$\Lambda(\underline{x}) = \frac{f(\underline{x} \mid H_1)}{f(\underline{x} \mid H_0)}$$

$$= \frac{(N+t+1)!}{N!\,(N+1)^{t+1} \prod_{i=1}^{N} x_i!} \overset{H_1}{\underset{H_0}{\gtrless}} \eta$$

Comments

1. We carefully chose our priors so that the integrals could be computed in closed form.
2. The "integrated LRT" is optimal only if we used the correct priors.
3. In general, the computationally convenient prior is not the correct prior, so we must
   (a) be content with a suboptimal detector, OR
   (b) resort to time consuming numerical integration techniques
4. Bayes factor has unknown distribution; must set threshold experimentally.

$\boxed{\text{Generalized Likelihood Ratio Tests (GLRTs)}}$

Consider two competing models

$$H_0 : \underline{X} \sim f(\underline{x} \mid \underline{\theta}_0)$$

$$H_1 : \underline{X} \sim f(\underline{x} \mid \underline{\theta}_1)$$

The models each have unknown parameters (not necessarily the same distributional family)

<u>Idea</u> Use the data to <u>estimate</u> the unknown parameters and <u>plug in</u> to the LRT

$$\boxed{\tilde{\Lambda}(\underline{x}) = \frac{f(\underline{x} \mid \hat{\underline{\theta}}_1)}{f(\underline{x} \mid \hat{\underline{\theta}}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta} \leftarrow \text{GLRT}$$

$\hat{\theta}_k = \hat{\theta}_k(\underline{x}) =$ data-based estimate

When estimating $\underline{\theta}$, we use the <u>maximum likelihood estimate</u> (MLE)

$$\underline{\hat{\theta}}_{ML} = \arg\max_{\underline{\theta}} f(\underline{x} \mid \underline{\theta})$$

In summary, the <u>GLRT</u> is

$$\boxed{\tilde{\Lambda}(\underline{x}) = \frac{\max_{\underline{\theta}_1} f(\underline{x} \mid H_1, \underline{\theta}_1)}{\max_{\underline{\theta}_0} f(\underline{x} \mid H_0, \underline{\theta}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta}$$

Notes

1. We maximize the numerator and denominator <u>separately</u>. We do <u>not</u> maximize their <u>ratio</u>.

2. It can be shown that under mild conditions the GLRT

$$\tilde{\Lambda}(\underline{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

is asymptotically (as $N \to \infty$) UMP among all decision rules that are <u>invariant</u> to the unknown parameters (i.e., that don't depend on the unknown parameters) See Kay, vol. II.

Ex | Geometric vs. Poisson

Geometric:

$$\hat{\theta}_{ML} | H_0 = \arg\max_{\theta} \left[\theta^N (1-\theta)^t \right], \quad t = \sum_{i=1}^{N} x_i$$

$$\frac{\partial}{\partial \theta} \left(\theta^N (1-\theta)^t \right) = N\theta^{N-1}(1-\theta)^t - \theta^N t (1-\theta)^{t-1}$$

set derivative to zero $\Rightarrow$

$$N\theta^{N-1}(1-\theta)^t = \theta^N t (1-\theta)^{t-1}$$

$$\Rightarrow N(1-\theta) = t\theta$$

$$\Rightarrow \hat{\theta}_{ML} | H_0 = \frac{1}{1+t/N} = \frac{1}{1+\frac{1}{N}\sum x_i}$$

Poisson

$$\hat{\theta}_{ML} | H_1 = \arg\max_{\theta} \left[e^{-N\theta} \frac{\theta^t}{\prod_{i=1}^{N} x_i!} \right]$$

$$\frac{\partial}{\partial \theta} \left(e^{-N\theta} \frac{\theta^t}{\prod x_i!} \right) = -N e^{-N\theta} \frac{\theta^t}{\prod x_i!} + e^{-N\theta} \frac{t\theta^{t-1}}{\prod x_i!}$$

set derivative to zero $\Rightarrow$

$$N e^{-N\theta} \frac{\theta^t}{\prod x_i!} = e^{-N\theta} \frac{t\theta^{t-1}}{\prod x_i!}$$

$$\Rightarrow N\theta = t$$

$$\Rightarrow \hat{\theta}_{ML} | H_1 = \frac{t}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

GLRT

$$\tilde{\Lambda}(\underline{x}) = \frac{\max_{\theta_1} f(\underline{x} \mid H_1, \theta_1)}{\max_{\theta_0} f(\underline{x} \mid H_0, \theta_0)}$$

$$= \frac{e^{-t}\left(t/N\right)^{t}}{\left(\prod_{i=1}^{N} x_i!\right)\left(\frac{1}{1+t/N}\right)^{N}\left(\frac{t/N}{1+t/N}\right)^{t}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

The GLR statistic does not involve the unknown parameters, but it is still difficult to set the threshold because the distribution of $\tilde{\Lambda}(\underline{x})$ is unknown.

The threshold must be set experimentally or through numerical simulations.

Exercise | Consider the detection problem

$$H_0: \quad X(n) = w(n)$$

$$n = 0, 1, \cdots, N-1$$

$$H_1: \quad X(n) = A + w(n)$$

where A is unknown and $w(n) \overset{iid}{\sim} N(0, \sigma^2)$, σ^2 known. Find the GLRT. Set the threshold to ensure a false alarm rate of α.

Solution

$$\hat{A}_{ML} = \arg\max \left[\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^{N} (x_i - A)^2 \right\} \right]$$

$$= \arg\min \left[\sum_{i=1}^{N} (x_i - A)^2 \right]$$

$$\frac{\partial}{\partial A}\left(\sum (x_i - A)^2 \right) = -2 \sum_{i=1}^{N} (x_i - A) = -0$$

$$\Rightarrow \hat{A}_{ML} = \frac{1}{N} \sum_{i=1}^{N} x_i \equiv t$$

$$\tilde{\Lambda}(\underline{x}) = \frac{\exp\left\{ \frac{-1}{2\sigma^2} \sum (x_i - \hat{A})^2 \right\}}{\exp\left\{ \frac{1}{2\sigma^2} \sum x_i^2 \right\}}$$

$$\Rightarrow \log \tilde{\Lambda}(\underline{x}) = \frac{1}{2\sigma^2}\left(-N\hat{A}^2 + 2\hat{A} \sum_{i=1}^{N} x_i \right)$$

$$= \frac{N}{2\sigma^2} \cdot \hat{A}^2 \underset{H_0}{\overset{H_1}{\gtrless}} \log(\eta)$$

$$\Rightarrow \boxed{|t| \gtrless \gamma \equiv \sqrt{\frac{2\sigma^2}{N} \log(\eta)}}$$

$$t = \hat{A} = \frac{1}{N} \sum x_i$$

Under H_0, $T \sim N(0, \frac{\sigma^2}{N})$ and therefore

$$P_F = P(|T| > \gamma \mid H_0)$$

$$= 2P(T > \gamma \mid H_0)$$

$$= 2Q\left(\frac{\gamma}{\sigma/\sqrt{N}}\right) = \alpha$$

$$\Rightarrow \quad \gamma = \frac{\sigma}{\sqrt{N}} Q^{-1}\left(\frac{\alpha}{2}\right)$$

Remarks/ 1. This is basically the same suboptimal test we studied last time.

2. A Bayes factor test with $A \sim N(0, \tau^2)$ leads to the same detector.

Unknown Signal Delay

Consider a binary test for the presence or absence of a signal. Assume that the signal waveform is known, but its time origin is not.

$$H_0: \quad x(n) = w(n)$$

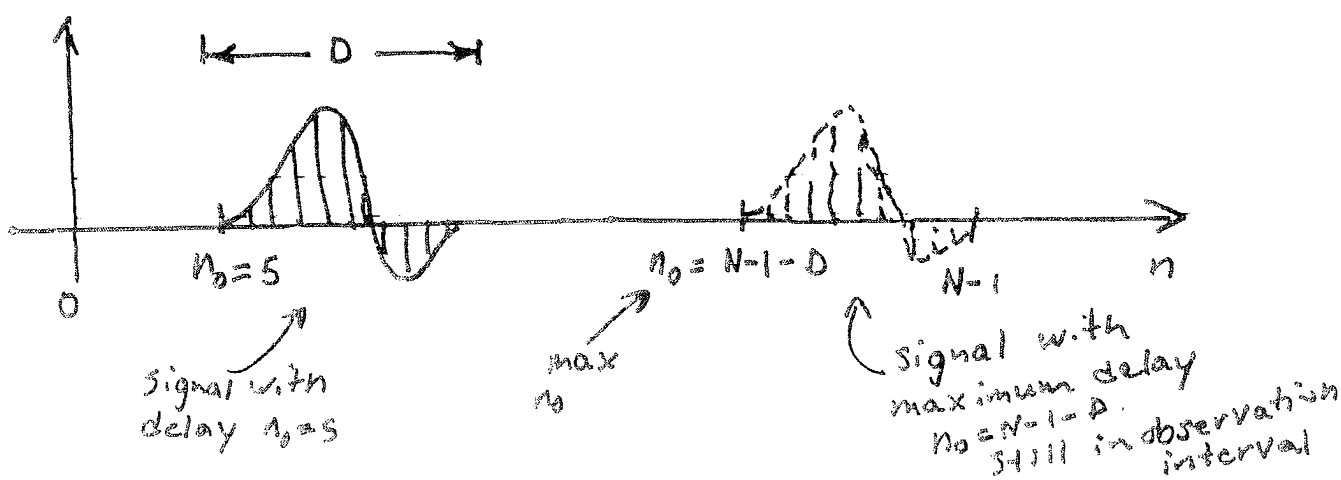
$$n = 0, 1, \ldots, N-1$$

$$H_1: \quad x(n) = s(n - n_0) + w(n)$$

n_0 is an unknown integer.

$\underline{w} \sim N(\underline{0}, \sigma^2 I)$.

We will also assume that for all possible values of n_0 the signal lies completely in the observation interval:



Under H_1

$$\underline{x} \sim f(\underline{x} \mid H_1, n_0)$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left[x(n) - s(n-n_0)\right]^2\right\}$$

If we form $\Lambda(\underline{x})$, the LR, it is easy to see that it depends on the unknown delay n_0 and a UMP test does not exist.

Therefore, we again will try the GLRT approach.

Maximizing $f(\underline{x} \mid H_1, n_0)$ with respect to n_0 is equivalent to maximizing

$$\sum_{n=0}^{N-1}\left[x(n)\, s(n-n_0) - \frac{1}{2} s^2(n-n_0)\right]$$

$$= \sum_{n=0}^{N-1}\left[x(n)\, s(n-n_0)\right] - \frac{1}{2}\|\underline{s}\|^2$$

The log GLRT is then

$$\log \tilde{\lambda}(\underline{x}) = \log \frac{\max_{n_0} f(\underline{x} \mid H_1, n_0)}{f(\underline{x} \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \log \eta$$

or equivalently

$$\max_{n_0} \sum_{n=n_0}^{n_0+D-1} x(n)s(n-n_0) \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\sigma^2 \log \eta + \frac{\|\underline{s}\|^2}{2}}_{\gamma}$$

Using the matched filter interpretation of the test statistic, the decision rule can be expressed as

$$\max_{n_0} \left[x(n) * s(D-1-n) \right] \Big|_{n=D-1+n_0} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

Hence, we simply compute each output of the matched filter (corresponding to each possible delay) and pick the largest value to evaluate the GLRT.

Note that the index at which the maximum occurs gives us the MLE of delay n_0.

$\Longrightarrow$ detection and estimation problems are solved simultaneously.

Can we set the threshold to achieve a specified P_F?

Summary

- UMP tests rarely exist so we need methods for designing suboptimal detectors
  - ☐ Bayesian approach
    - → Bayes factor
  - ☐ Classical approach
    - → GLRT
- In practice GLRT is more widely used since integrated likelihoods are often difficult to compute
- Application: signal delay estimation / detection via matched filtering.