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 $\Theta$:

1. $\Theta$ is itself a random quantity (Bayesian approach)
2. $\Theta$ is unknown, but fixed (Classical approach)

Bayes Factors

Consider

$$H_0 : X \sim f(x | \Theta_0)$$
$$H_1 : X \sim f(x | \Theta_1)$$

where $\Theta_0$ and $\Theta_1$ are unknown.

Assume $\Theta_0$ and $\Theta_1$ are realizations of the prior distributions

$$f(\Theta_k | H_k), \quad k = 0, 1$$
Then

\[ f(x | H_k) = \int f(x, \theta_k | H_k) \, d\theta_k \]

\[ = \int f(x | H_k, \theta_k) \, f(\theta_k | H_k) \, d\theta_k \]

Thus the LR statistic is

\[ \Lambda(x) = \frac{f(x | H_1)}{f(x | H_0)} \]

\[ = \frac{\int f(x | H_1, \theta_1) \, f(\theta_1 | H_1) \, d\theta_1}{\int f(x | H_0, \theta_0) \, f(\theta_0 | H_0) \, d\theta_0} \]

This "integrated likelihood ratio" is called the Bayes factor for testing \( H_1 \) vs. \( H_0 \).
Example: Geometric vs. Poisson

\( H_0 : \quad x_1, \ldots, x_N \text{ iid geometric} \)

\[ x_i \sim \theta_0 (1-\theta_0)^{x_i}, \quad 0 \leq \theta_0 \leq 1 \]

\[ f(x | H_0, \theta_0) = \theta_0^N (1 - \theta_0)^{\sum_{i=1}^{N} x_i} \]

\( H_1 : \quad x_1, \ldots, x_N \text{ iid Poisson} \)

\[ x_i \sim e^{-\theta_1} \frac{\theta_1^{x_i}}{x_i!}, \quad \theta_1 > 0 \]

\[ f(x | H_1, \theta_1) = e^{-N\theta_1} \frac{\theta_1^{\sum_{i=1}^{N} x_i}}{N \prod_{i=1}^{N} x_i!} \]

Can we apply Karlin-Rubin?
Taking a Bayesian approach, let's assume prior distributions for $\Theta_0$ and $\Theta_1$:

**Modeling uncertainty in $\Theta_0$**

Under $H_0$: $0 \leq \Theta_0 \leq 1$

$$f(\Theta_0) = I_{[0,1]}(\Theta_0)$$

uniform: no preference for any value

Under $H_1$: $\Theta_1 > 0$

$$f(\Theta_1) = e^{-\Theta_1}$$

exponential: favors smaller values

**Integrated Likelihoods**

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

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

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

\[ \Lambda(x) = \frac{f(x \mid H_1)}{f(x \mid H_0)} \]

\[ = \frac{(N+t+1)!}{N! (N+1)^{t+1}} \prod_{i=1}^{N} \frac{x_i!}{\eta_i} \]

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.
Generalized Likelihood Ratio Tests (GLRTs)

Consider two competing models

\[ H_0 \colon x \sim f(x \mid \theta_0) \]
\[ H_1 \colon x \sim f(x \mid \theta_1) \]

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

**Ideal**: Use the data to estimate the unknown parameters and plug in to the LRT

\[ \Lambda(x) = \frac{f(x \mid \hat{\theta}_1)}{f(x \mid \hat{\theta}_0)} \]

\[ H_1 \quad \text{if} \quad \Lambda(x) > \eta \quad \leftarrow \text{GLRT} \]

\[ \hat{\theta}_k = \hat{\theta}_k(x) = \text{data-based estimate} \]
When estimating $\Theta$, we use the maximum likelihood estimate (MLE)

$$\hat{\Theta}_{ML} = \arg \max_{\Theta} f(x | \Theta)$$

In summary, the GLRT is

$$\Lambda(x) = \max_{\Theta_0} f(x | H_0, \Theta_0)$$
$$\Lambda(x) = \max_{H_1} \frac{f(x | H_1, \Theta_1)}{f(x | H_0, \Theta_0)} \quad H_1 \gg H_0$$
1. We maximize the numerator and denominator separately. We do not maximize their ratio.

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

\[ \Lambda(x) \overset{H_1}{\underset{H_0}{\gtrless}} \eta \]

is asymptotically (as \( N \to \infty \)) UMP among all decision rules that are invariant to the unknown parameters (i.e., that don't depend on the unknown parameters) See Kay, vol. II.
\textbf{Ex. Geometric vs. Poisson}

\textbf{Geometric:}
\[ \hat{\theta}_{ML} | H_0 = \arg \max_{\theta} \left[ \theta^N (1-\theta)^t \right], \quad t = \sum_{i=1}^{n} x_i \]
\[ \frac{\partial}{\partial \theta} \left( \theta^N (1-\theta)^t \right) = N\theta^{N-1}(1-\theta)^t - \theta^t (1-\theta)^{t-1} \]
set derivative to zero \( \Rightarrow \)
\[ N\theta^{N-1}(1-\theta)^t = \theta^t (1-\theta)^{t-1} \]
\[ \Rightarrow N(1-\theta) = t\theta \]
\[ \Rightarrow \hat{\theta}_{ML} | H_0 = \frac{1}{1 + t/N} = \frac{1}{1 + \frac{1}{N} \sum x_i} \]

\textbf{Poisson}
\[ \hat{\theta}_{ML} | H_1 = \arg \max_{\theta} \left[ e^{-N\theta} \frac{\theta^t}{\prod_{i=1}^{n} x_i!} \right] \]
\[ \frac{\partial}{\partial \theta} \left( e^{-N\theta} \frac{\theta^t}{\prod_{i=1}^{n} x_i!} \right) = -Ne^{-N\theta} \frac{\theta^t}{\prod_{i=1}^{n} x_i!} + e^{-N\theta} \frac{t \theta^{t-1}}{\prod_{i=1}^{n} x_i!} \]
set derivative to zero \( \Rightarrow \)
\[ Ne^{-N\theta} \frac{\theta^t}{\prod_{i=1}^{n} x_i!} = e^{-N\theta} \frac{t \theta^{t-1}}{\prod_{i=1}^{n} x_i!} \]
\[ \Rightarrow N\theta = t \]
\[ \Rightarrow \hat{\theta}_{ML} | H_1 = \frac{t}{N} = \frac{1}{N} \sum_{i=1}^{n} x_i \]
\[ \tilde{\Lambda}(\mathbf{x}) = \frac{\max_{\Theta_i} f(\mathbf{x} | H_i, \Theta_i)}{\max_{\Theta_0} f(\mathbf{x} | H_0, \Theta_0)} \]

\[ = \frac{e^{-\frac{t}{N}} \left( \frac{t}{N} \right)^t}{\left( \prod_{i=1}^{N} x_i \right) \left( \frac{1}{1+t/N} \right)^N \left( \frac{t/N}{1+t/N} \right)^t} \quad \frac{H_1}{H_0} \geq \eta \]

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

The threshold must be set experimentally or through numerical simulations.
Exercise  
Consider the detection problem

\[ H_0 : \quad x(n) = w(n) \quad n = 0, 1, \ldots, N-1 \]
\[ H_1 : \quad x(n) = A + w(n) \]

where \( A \) is unknown and \( w(n) \) iid \( N(0, \sigma^2) \), \( \sigma^2 \) known. Find the GLRT. Set the threshold to ensure a false alarm rate of \( \alpha \).
Solution

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

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

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

\[ \Rightarrow \hat{A}_{\text{ML}} = \frac{1}{N} \sum_{i=1}^{N} x_i = t \]

\[ \tilde{\lambda}(x) = \frac{\exp \left\{ \frac{-1}{2\sigma^2} \sum (x_i - \hat{A})^2 \right\}}{\exp \left\{ \frac{1}{2\sigma^2} \sum x_i^2 \right\}} \]

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

\[ = \frac{N}{2\sigma^2} \hat{A}^2 \quad \Rightarrow \quad \log \left( \frac{H_1}{H_0} \right) \]

\[ \Rightarrow \quad t \leq \hat{\sigma} = \sqrt{\frac{2\sigma^2}{N} \log (\eta)} \]

\[ t = \hat{A} = \frac{1}{N} \sum x_i \]
Under $H_0$, $T \sim N(0, \frac{\sigma^2}{N})$ and therefore

$$P_F = P( |T| > \delta | H_0)$$

$$= 2 P( T > \delta | H_0)$$

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

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

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.
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.

\[ H_0 : x(n) = w(n) \quad n = 0, 1, \ldots, N-1 \]

\[ H_1 : x(n) = s(n-n_0) + w(n) \]

\( n_0 \) is an unknown integer.

\( w \sim N(0, \sigma^2 I) \).

We will also assume that for all possible values of \( n_0 \) the signal lies completely in the observation interval.

![Signal with delay](image)
Under $H_1$,

$$x \sim f(x \mid H_1, n_0)$$

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

If we form $\Lambda(x)$, 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 $f(x \mid H_1, n_0)$ with respect to $n_0$ is equivalent to maximizing

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

$$= \sum_{n=0}^{N-1} \left[ x(n) s(n-n_0) \right] - \frac{1}{2} \| s \|^2$$
The log GLRT is then

\[
\log \hat{\lambda}(x) = \log \max_{n_0} \frac{f(x | H_1, n_0)}{f(x | H_0)} \quad H_1 \overset{\text{\(\succ\)}}{\sim} \log n \quad H_0
\]

or equivalently

\[
\max_{n_0} \sum_{n=n_0}^{n_0+D-1} x(n) s(n-n_0) \quad H_1 \overset{\text{\(\succ\)}}{\sim} \sigma^2 \log n + \frac{\|s\|^2}{2}
\]

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

\[
\max_{n_0} \left[ (\hat{x}(n) \ast s(D-1-n)) \right]_{n=n_0+D-1} \quad H_1 \overset{\text{\(\succ\)}}{\sim} \gamma \quad H_0
\]

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 $n_0$.

⇒ detection and estimation problems are solved simultaneously.

Can we set the threshold to achieve a specified $P_e$?
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.