Outline:

- Neyman-Pearson test for simple binary hypotheses, receiver operating characteristic (ROC).

- An introduction to classical composite hypothesis testing.

Reading:

- Chapter 3 in Kay-II,

- (part of) Chapter 5 in Levy.
In binary hypothesis testing, we wish to identify which hypothesis is true (i.e. make the appropriate decision):

\[ \mathcal{H}_0 \ : \ \theta \in \text{sp}_\Theta(0) \quad \text{null hypothesis} \quad \text{versus} \]
\[ \mathcal{H}_1 \ : \ \theta \in \text{sp}_\Theta(1) \quad \text{alternative hypothesis} \]

where

\[ \text{sp}_\Theta(0) \cup \text{sp}_\Theta(1) = \text{sp}_\Theta, \ \text{sp}_\Theta(0) \cap \text{sp}_\Theta(1) = \emptyset. \]

Recall that a binary decision rule \( \phi(x) \) maps data space \( \mathcal{X} \) to \( \{0, 1\} \):

\[
\phi(x) = \begin{cases} 
1, & \text{decide } \mathcal{H}_1, \\
0, & \text{decide } \mathcal{H}_0.
\end{cases}
\]

which partitions the data space \( \mathcal{X} \) [i.e. the support of \( f_x |_\Theta(x | \theta) \)] into two regions:

\[ \mathcal{X}_0 = \{x : \phi(x) = 0\} \quad \text{and} \quad \mathcal{X}_1 = \{x : \phi(x) = 1\}. \]
Recall the probabilities of false alarm and miss:

\[
P_{FA}(\phi(X), \theta) = \mathbb{E}_{X|\Theta}[\phi(X) | \theta] = \int_{x_1} f_{X|\Theta}(x | \theta) \, dx \quad \text{for } \theta \in \text{sp}\Theta(0) \tag{1}
\]

\[
P_{M}(\phi(X), \theta) = \mathbb{E}_{X|\Theta}[1 - \phi(X) | \theta] = 1 - \int_{x_1} f_{X|\Theta}(x | \theta) \, dx
\]

\[
= \int_{x_0} f_{X|\Theta}(x | \theta) \, dx \quad \text{for } \theta \text{ in } \text{sp}\Theta(1) \tag{2}
\]

and the probability of detection (correctly deciding \(H_1\)):

\[
P_{D}(\phi(X), \theta) = \mathbb{E}_{X|\Theta}[\phi(X) | \theta] = \int_{x_1} f_{X|\Theta}(x | \theta) \, dx \quad \text{for } \theta \text{ in } \text{sp}\Theta(1).
\]

For simple hypotheses, \(\text{sp}\Theta(0) = \{\theta_0\}, \text{sp}\Theta(1) = \{\theta_1\}, \) and \(\text{sp}\Theta = \{\theta_0, \theta_1\},\) the above expressions simplify, as shown in the following.
Probabilities of False Alarm ($P_{FA}$) and Detection ($P_D$) for Simple Hypotheses

\[ P_{FA}(\phi(X), \theta_0) = \int_{X_1} f_X|_{\Theta}(x|\theta_0) \, dx \]
\[ = \Pr_{X|\Theta}\{\text{test statistic (ts)} > \tau \mid \theta_0 \} \quad (3) \]

\[ P_D(\phi(X), \theta_1) = \int_{X_1} f_X|_{\Theta}(x|\theta_1) \, dx \]
\[ = \Pr_{X|\Theta}\{\text{ts} > \tau \mid \theta_1 \}. \quad (4) \]

Comments:

(i) As the region $X_1$ shrinks (i.e. $\tau \nearrow \infty$), both of the above
probabilities shrink towards zero.

**(ii)** As the region $\mathcal{X}_1$ grows (i.e. $\tau \downarrow 0$), both probabilities grow towards unity.

**(iii)** Observations (i) and (ii) do not imply equality between $P_{FA}$ and $P_D$; in most cases, as $\mathcal{X}_1$ grows, $P_D$ grows more rapidly than $P_{FA}$ (i.e. we better be right more often than we are wrong).

**(iv)** However, the perfect case where our rule is always right and never wrong ($P_D = 1$ and $P_{FA} = 0$) cannot occur when the conditional pdfs/pmfs $f_{X \mid \Theta}(x \mid \theta_0)$ and $f_{X \mid \Theta}(x \mid \theta_1)$ overlap.

**(v)** Thus, to increase the detection probability $P_D$, we must also allow for the false-alarm probability $P_{FA}$ to increase. This behavior

- represents the fundamental tradeoff in hypothesis testing and detection theory and
- motivates us to introduce a (classical) approach to testing simple hypotheses, pioneered by Neyman and Pearson, to be discussed next.

Receiver Operating Characteristic (ROC) allows us to visualize the realm of achievable $P_{FA}(\phi(X), \theta_0)$ and $P_D(\phi(X), \theta_1)$. 
A point \((P_{FA}, P_D)\) is in the shaded region if we can find a rule \(\phi(X)\) such that \(P_{FA}(\phi(X), \theta_0) = P_{FA}\) and \(P_D(\phi(X), \theta_1) = P_D\).
Bayesian tests are criticized because they require specification of prior distribution (pmf or, in the composite-testing case, pdf) and the cost-function parameters $L(i \mid j)$.

An alternative classical solution for simple hypotheses is developed by Neyman and Pearson.

Select the decision rule $\phi(X)$ that maximizes $P_D(\phi(X), \theta_1)$ while ensuring that the probability of false alarm $P_{FA}(\phi(X), \theta_0)$ is less than or equal to a specified level $\alpha$.

Setup:

- **Simple hypothesis** testing:
  \[ H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1. \]

- Parametric data models $f_X(\mathbf{x} \mid \theta_0)$, $f_X(\mathbf{x} \mid \theta_1)$.

- No prior pdf/pmf on $\Theta$ is available.
Define the set of all rules $\phi(X)$ whose probability of false alarm is less than or equal to a specified level $\alpha$:

$$D_\alpha = \{ \phi(X) \mid P_{FA}(\phi(X), \theta_0) \leq \alpha \}$$

see also (3).

A Neyman-Pearson test $\phi_{NP}(x)$ solves the constrained optimization problem:

$$\phi_{NP}(x) = \arg \max_{\phi(x) \in D_\alpha} P_D(\phi(x), \theta_1).$$

We apply Lagrange multipliers to solve this optimization problem; consider the Lagrangian:

$$L(\phi(x), \lambda) = P_D(\phi(x), \theta_1) + \lambda [\alpha - P_{FA}(\phi(x), \theta_0)]$$

with $\lambda \geq 0$. A decision rule $\phi(x)$ will be optimal if it maximizes $L(\phi(x), \lambda)$ and satisfies the Karush-Kuhn-Tucker (KKT) condition:

$$\lambda [\alpha - P_{FA}(\phi(x), \theta_0)] = 0. \quad (5)$$

Upon using (3) and (4), the Lagrangian can be written as

$$L(\phi(x), \lambda) = \lambda \alpha + \int_{\chi_1} [f_{X \mid \Theta}(x \mid \theta_1) - \lambda f_{X \mid \Theta}(x \mid \theta_0)] dx.$$
Consider maximizing $L(\phi(x), \lambda)$ with respect to $\phi(x)$ for a given $\lambda$. Then, $\phi(x)$ needs to satisfy

$$
\phi_\lambda(x) = \begin{cases} 
1, & \Lambda(x) > \lambda \\
0 \text{ or } 1, & \Lambda(x) = \lambda \\
0, & \Lambda(x) < \lambda 
\end{cases}
$$

(6)

where

$$
\Lambda(x) = \frac{f_{X|\Theta}(x|\theta_1)}{f_{X|\Theta}(x|\theta_0)}
$$

is the likelihood ratio. The values $x$ that satisfy $\Lambda(x) = \lambda$ can be allocated to either $X_1$ or $X_0$. To completely specify the optimal test, we need to select

- a $\lambda$ such that the KKT condition (5) holds and
- an allocation rule for those $x$ that satisfy $\Lambda(x) = \lambda$.

Now, consider two versions of (6) for a fixed threshold $\lambda$:

$$
\phi_{U,\lambda}(x) = \begin{cases} 
1, & \Lambda(x) > \lambda \\
1, & \Lambda(x) = \lambda \\
0, & \Lambda(x) < \lambda 
\end{cases}
$$

and

$$
\phi_{L,\lambda}(x) = \begin{cases} 
1, & \Lambda(x) > \lambda \\
0, & \Lambda(x) = \lambda \\
0, & \Lambda(x) < \lambda 
\end{cases}
$$
In the first case, all observations \( x \) for which \( \Lambda(x) = \lambda \) are allocated to \( \mathcal{X}_1 \); in the second case, these observations are allocated to \( \mathcal{X}_0 \).

Consider the cumulative distribution function (cdf) of \( \Lambda(X) = \Lambda \) under \( \mathcal{H}_0 \):

\[
F_{\Lambda|\Theta}(l|\theta_0) = \Pr_{\Lambda|\Theta}\{\Lambda \leq l | \theta_0\}.
\]

Define

\[
f_0 = F_{\Lambda|\Theta}(0|\theta_0) = \Pr_{\Lambda|\Theta}\{\Lambda \leq 0 | \theta_0\}.
\]

Recall that cdf \( F_{\Lambda|\Theta}(l|\theta_0) \) must be nondecreasing and right-continuous, but may have discontinuities.
Consider three cases, depending on $\alpha$:

(i) When

$$1 - \alpha < f_0 \quad \text{i.e.} \quad 1 - f_0 < \alpha$$  \hspace{1cm} (7)

we select the threshold $\lambda = 0$ and apply the rule

$$\phi_{L,0}(x) = \begin{cases} 
1, & \Lambda(x) > 0 \\
0, & \Lambda(x) = 0 
\end{cases} .$$  \hspace{1cm} (8)

In this case, KKT condition (5) holds and, therefore, the test (8) is optimal; its probability of false alarm is

$$P_{FA}(\phi_{L,0}(x), \theta_0) = 1 - f_0 < \alpha \quad \text{see (7)}.$$

An example of this case corresponds to $\lambda_1 = 0$ and $1 - \alpha_1$ in the above figure.

(ii) Suppose that

$$1 - \alpha \geq f_0 \quad \text{i.e.} \quad 1 - f_0 \geq \alpha$$  \hspace{1cm} (9)

and there exists a $\lambda$ such that

$$F_{\Lambda|\Theta}(\lambda | \theta_0) = 1 - \alpha.$$  \hspace{1cm} (10)

Then, by selecting this $\lambda$ as the threshold and using

$$\phi_{L,\lambda}(x) = \begin{cases} 
1, & \Lambda(x) > \lambda \\
0, & \Lambda(x) \leq \lambda 
\end{cases}$$  \hspace{1cm} (11)
we obtain a test with false-alarm probability

\[ P_{FA}(\phi_L, \lambda(x), \theta_0) = 1 - F_{\Lambda | \Theta}(\lambda | \theta_0) = \alpha \]  

see (9)

the KKT condition (5) holds, and the test (10) is optimal. An example of this case corresponds to \( \lambda_2 \) and \( 1 - \alpha_2 \) in the above figure.

(iii) Suppose that

\[ 1 - \alpha \geq f_0 \quad \text{i.e.} \quad 1 - f_0 \geq \alpha \]

as in (ii), but cdf \( F_{\Lambda | \Theta}(l | \theta_0) \) has a discontinuity point \( \lambda > 0 \) such that

\[ F_{\Lambda | \Theta}(\lambda_- | \theta_0) < 1 - \alpha < F_{\Lambda | \Theta}(\lambda_+ | \theta_0) \]

where \( F_{\Lambda | \Theta}(\lambda_- | \theta_0) \) and \( F_{\Lambda | \Theta}(\lambda_+ | \theta_0) \) denote the left and right limits of \( F_{\Lambda | \Theta}(\lambda | \theta_0) \) at \( l = \lambda \). If this case happens in practice, we can try to avoid the problem by changing our specified \( \alpha \), which is anyway not God-given, but chosen rather arbitrarily. We should pick a value of \( \alpha \) that satisfies the KKT condition.

Suppose that we are not allowed to change \( \alpha \); this gives us a chance to practice some basic probability. First, note that
• $\phi_{L,\lambda}(\mathbf{x})$ has false-alarm probability

$$P_{FA}(\phi_{L,\lambda}(\mathbf{x}), \theta_0) = 1 - F_{\Lambda|\Theta}(\lambda_+ | \theta_0) < \alpha,$$

• $\phi_{U}(\mathbf{x}, \lambda)$ has false-alarm probability

$$P_{FA}\{\phi_{U,\lambda}(\mathbf{x}), \theta_0\} = 1 - F_{\Lambda|\Theta}(\lambda_- | \theta_0) > \alpha$$

and KKT optimality condition (5) requires that $P_{FA}(\phi_{\lambda}(\mathbf{x}), \theta_0) = \alpha$. We focus on the tests of the form (6) and construct the optimal test via randomization.

Define the probability

$$p = \frac{\alpha - P_{FA}(\phi_{L,\lambda}(\mathbf{x}), \theta_0)}{P_{FA}(\phi_{U,\lambda}(\mathbf{x}), \theta_0) - P_{FA}(\phi_{L,\lambda}(\mathbf{x}), \theta_0)}$$

which clearly satisfies $0 < p < 1$. 
Select $\phi_{U,\lambda}(x)$ with probability $p$ and $\phi_{L,\lambda}(x)$ with probability $1 - p$. This test indeed has the form (6); its probability of false alarm is

\[
P_{FA}(\phi_{\lambda}(x), \theta_0) = P_{FA}(\phi_{L,\lambda}(x), \theta_0) + p \left[ P_{FA}(\phi_{U,\lambda}(x), \theta_0) - P_{FA}(\phi_{L,\lambda}(x), \theta_0) \right] = \alpha.
\]

Since KKT condition (5) is satisfied, the randomized test

\[
\phi_{\lambda}(x) = \begin{cases} 
1 \text{ w.p. } p & \text{and } 0 \text{ w.p. } 1 - p, \\
1, & \Lambda(x) > \lambda \\
0, & \Lambda(x) = \lambda \\
0, & \Lambda(x) < \lambda 
\end{cases}
\]

is optimal.
Based on the Neyman-Pearson theory, if we set $P_{FA} = \alpha$, then the test that maximizes $P_D$ must be a likelihood-ratio test of the form (6). Thus, the ROC curve separating achievable and non-achievable pairs ($P_{FA}, P_D$) corresponds to the family of likelihood-ratio tests.

For simplicity, we focus here on the case where the likelihood ratio is a continuous random variable given $\theta$. First, note that,
for the likelihood-ratio test,

\[ P_{\text{FA}}(\tau) = \int_{X_1} f_{X|\Theta}(x|\theta_0) \, dx \]

\[ = \Pr_{X|\Theta}\{\Lambda(X) > \tau | \theta_0\} = \int_{\tau}^{+\infty} f_{\Lambda|\Theta}(l|\theta_0) \, dl \tag{12} \]

\[ P_{\text{D}}(\tau) = \int_{X_1} f_{X|\Theta}(x|\theta_1) \, dx \]

\[ = \Pr_{X|\Theta}\{\Lambda(X) > \tau | \theta_1\} = \int_{\tau}^{+\infty} f_{\Lambda|\Theta}(l|\theta_1) \, dl \tag{13} \]

where \( \tau \) denotes the threshold. Under the continuity assumption for the likelihood ratio, as we vary \( \tau \) between 0 and \(+\infty\), the point \((P_{\text{FA}}(\phi(X), \theta_0), P_{\text{D}}(\phi(X), \theta_1))\) moves continuously along the ROC curve. If we set \( \tau = 0 \), we always select \( H_1 \) and, therefore,

\[ P_{\text{FA}}(0) = P_{\text{D}}(0) = 1. \]

Conversely, if we set \( \tau = +\infty \), we always select \( H_0 \) and, therefore,

\[ P_{\text{FA}}(+\infty) = P_{\text{D}}(+\infty) = 0. \]

In summary,

**ROC Property 1.** If the likelihood ratio is a continuous random variable given \( \theta \), the points \((0, 0)\) and \((1, 1)\) belong to ROC.
Now, differentiate (12) and (13) with respect to $\tau$:

$$\frac{dP_D(\tau)}{d\tau} = -f_{\Lambda|\Theta}(\tau | \theta_1)$$

$$\frac{dP_D(\tau)}{d\tau} = -f_{\Lambda|\Theta}(\tau | \theta_0)$$

implying

$$\frac{dP_D(\tau)}{dP_{FA}(\tau)} = \frac{f_{\Lambda|\Theta}(\tau | \theta_1)}{f_{\Lambda|\Theta}(\tau | \theta_0)} = \tau.$$

In summary,

**ROC Property 2.** *If the likelihood ratio is a continuous random variable given $\theta$, the slope of ROC at point $(P_{FA}(\tau), P_D(\tau))$ is equal to the threshold $\tau$ of the corresponding likelihood-ratio test.*

In particular, this result implies that the slope of ROC is

- $\tau = +\infty$ at $(0, 0)$ and
- $\tau = 0$ at $(1, 1)$.

**ROC Property 3.** *The domain of achievable pairs $(P_{FA}, P_D)$ is convex and the ROC curve is concave. This property holds in general, including the case where the likelihood ratio is a mixed or discrete random variable given $\theta$.*

HW: Prove ROC Property 3.
ROC Property 4. All points on ROC curve satisfy

\[ P_D(\tau) \geq P_{FA}(\tau). \]

This property holds in general, including the case where the likelihood ratio is a mixed or discrete random variable given \( \theta \).
Simple hypotheses: the space of the parameter $\mu$ and its partitions are

$$\mathcal{sp}_\mu = \{\mu_0, \mu_1\}, \quad \mathcal{sp}_\mu(0) = \{\mu_0\}, \quad \mathcal{sp}_\mu(1) = \{\mu_1\}.$$  

The measurement vector $X$ given $\mu$ is modeled using

$$f_{X \mid \mu}(x \mid \mu) = \mathcal{N}(x \mid \mu, C)$$

$$= \frac{1}{\sqrt{\mid 2\pi C \mid}} \exp[-\frac{1}{2}(x - \mu)^T C^{-1} (x - \mu)]$$

where $C$ is a known positive definite covariance matrix. Our likelihood-ratio test is

$$\Lambda(x) = \frac{f_{X \mid \mu}(x \mid \mu_1)}{f_{X \mid \mu}(x \mid \mu_0)}$$

$$= \frac{\exp[-\frac{1}{2}(x - \mu_1)^T C^{-1} (x - \mu_1)]}{\exp[-\frac{1}{2}(x - \mu_0)^T C^{-1} (x - \mu_0)]} \overset{\mathcal{H}_1}{\geq} \tau.$$
Therefore,

\[-\frac{1}{2} (x - \mu_1)^T C^{-1} (x - \mu_1) + \frac{1}{2} (x - \mu_0)^T C^{-1} (x - \mu_0) \gtrless H_1 \ln \tau\]

i.e.

\[(\mu_1 - \mu_0)^T C^{-1} [x - \frac{1}{2} (\mu_0 + \mu_1)] \gtrless \ln \tau.\]

and, finally,

\[T(x) = s^T C^{-1} x \overset{H_1}{\gtrless} \ln \tau + \frac{1}{2} (\mu_1 - \mu_0)^T C^{-1} (\mu_1 + \mu_0) \overset{\triangle}{=} \gamma\]

where we have defined

\[s \overset{\triangle}{=} \mu_1 - \mu_0.\]

**False-alarm and detection/miss probabilities.** Given \(\mu\), \(T(x)\) is a linear combination of Gaussian random variables, implying that it is also Gaussian, with mean and variance:

\[
E_{x_{|\mu}}[T(X)|\mu] = s^T C^{-1} \mu
\]

\[
\text{var}_{x_{|\mu}}[T(X)|\mu] = s^T C^{-1} s \quad (\text{not a function of } \mu).
\]
Now,

\[
P_{\text{FA}} = \Pr_{X | \mu} \{ T(X) > \gamma | \mu_0 \} \\
= \Pr_{X | \mu} \left\{ \frac{T(X) - s^T C^{-1} \mu_0}{\sqrt{s^T C^{-1} s}} > \frac{\gamma - s^T C^{-1} \mu_0}{\sqrt{s^T C^{-1} s}} \right\} | \mu_0 \}
\]

\[
= Q \left( \frac{\gamma - s^T C^{-1} \mu_0}{\sqrt{s^T C^{-1} s}} \right) \tag{14}
\]

and

\[
P_D = 1 - P_M = \Pr_{X | \mu} \{ T(X) > \gamma | \mu_1 \} \\
= \Pr_{X | \mu} \left\{ \frac{T(X) - s^T C^{-1} \mu_1}{\sqrt{s^T C^{-1} s}} > \frac{\gamma - s^T C^{-1} \mu_1}{\sqrt{s^T C^{-1} s}} \right\} | \mu_1 \}
\]

\[
= Q \left( \frac{\gamma - s^T C^{-1} \mu_1}{\sqrt{s^T C^{-1} s}} \right).
\]

We use (14) to obtain a \(\gamma\) that satisfies the specified \(P_{\text{FA}}\):

\[
\frac{\gamma}{\sqrt{s^T C^{-1} s}} = Q^{-1}(P_{\text{FA}}) + \frac{s^T C^{-1} \mu_0}{\sqrt{s^T C^{-1} s}}
\]
implying

\[ P_D = Q \left( Q^{-1}(P_{FA}) - \sqrt{s^T C^{-1} s} \right) \]

\[ = Q \left( Q^{-1}(P_{FA}) - d \right) \]  \hspace{1cm} (15)

Here,

\[ d = \sqrt{s^T C^{-1} s} = \sqrt{(\mu_1 - \mu_0)^T C^{-1} (\mu_1 - \mu_0)} \]

is the deflection coefficient.
Decentralized Detection for Simple Hypotheses

Consider a decentralized detection scenario depicted by

Assumptions:

- The observations $X[n]$, $n = 0, 1, \ldots, N - 1$ made at $N$ spatially distributed sensors (nodes) follow the same marginal probabilistic model:

$$f_{X|\Theta}(x[n] | \theta)$$

and are conditionally independent given $\Theta = \theta$, which may not always be reasonable, but leads to an easy solution.
• We wish to test:

\[ H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1. \]

• Each node \( n \) makes a **hard local decision** \( d[n] \) based on its local observation \( x[n] \) and sends it to the **headquarters** (fusion center), which collects all the local decisions and makes the final **global decision** \( H_0 \) versus \( H_1 \). This structure is clearly suboptimal: it is easy to construct a better decision strategy in which each node sends its (quantized, in practice) likelihood ratio to the fusion center, rather than the decision only. However, such a strategy would have a higher communication (energy) cost.

The false-alarm and detection probabilities of each node’s local decision rules can be computed using (16). Suppose that we have obtained them for each \( n \):

\[ P_{FA,n}, P_{D,n}, \quad n = 0, 1, \ldots, N - 1. \]

We now discuss the decentralized detection problem. Note that

\[
p_{D(n)} | \Theta(d_n | \theta_1) = P_{D,n}^{d_n} (1 - P_{D,n})^{1-d_n}
\]

Bernoulli pmf
and, similarly,

\[ p_{D(n)} | \Theta(d_n | \theta_0) = P_{d,n}^{d_n} (1 - P_{FA,n})^{1-d_n} \]  

where \( P_{FA,n} \) is the \( n \)th sensor’s local detection false-alarm probability. Now,

\[
\ln \Lambda(d) = \sum_{n=1}^{N} \ln \left[ \frac{p_{D(n)} | \Theta(d_n | \theta_1)}{p_{D(n)} | \Theta(d_n | \theta_0)} \right]
\]

\[
= \sum_{n=1}^{N} \ln \left[ \frac{P_{D,n}^{d_n} (1 - P_{D,n})^{1-d_n}}{P_{FA,n}^{d_n} (1 - P_{FA,n})^{1-d_n}} \right] \geq \ln \tau.
\]

To be able to further simplify the above expression, we now focus on the case where all sensors have identical performance:

\[ P_{D,n} = P_{D}, \quad P_{FA,n} = P_{FA} \]

i.e. all local decision thresholds at the nodes are identical. Define the number of sensors deciding locally to support \( H_1 \):

\[ u_1 = \sum_{n=0}^{N-1} d[n]. \]
Then, the log-likelihood ratio becomes

$$
\log \Lambda(d) = u_1 \log \left( \frac{P_D}{P_{FA}} \right) + (N - u_1) \log \left( \frac{1 - P_D}{1 - P_{FA}} \right) \overset{\mathcal{H}_1}{\geq} \log \tau
$$

or

$$
u_1 \log \left[ \frac{P_D \cdot (1 - P_{FA})}{P_{FA} \cdot (1 - P_D)} \right] \overset{\mathcal{H}_1}{\geq} \log \tau + N \log \left( \frac{1 - P_{FA}}{1 - P_D} \right). \quad (17)
$$

Clearly, each node’s local decision $d_n$ is meaningful only if $P_D > P_{FA}$, which implies

$$
\frac{P_D \cdot (1 - P_{FA})}{P_{FA} \cdot (1 - P_D)} > 1
$$

the logarithm of which is therefore positive, and the decision rule (17) further simplifies to

$$
u_1 \overset{\mathcal{H}_1}{\geq} \tau'.
$$

The Neyman-Person performance analysis of this detector is easy: the random variable $U_1$ is binomial given $\theta$ (i.e. conditional on the hypothesis) and, therefore,

$$
\Pr_{U_1 \mid \Theta} \{ U_1 = u_1 \mid \theta \} = \binom{N}{u_1} p^{u_1} (1 - p)^{N-u_1}
$$
where \( p = P_{\text{FA}} \) under \( \mathcal{H}_0 \) and \( p = P_D \) under \( \mathcal{H}_1 \). Hence, the “global” false-alarm probability is

\[
P_{\text{FA,global}} = \Pr_{U_1 \mid \Theta} \{ U_1 > \tau' \mid \theta_0 \}
\]

\[
= \sum_{u_1 = \lceil \tau' \rceil}^{N} \binom{N}{u_1} \cdot P_{\text{FA}}^{u_1} \cdot (1 - P_{\text{FA}})^{N-u_1}.
\]
An Introduction to Classical Composite Hypothesis Testing

First, recall that, in composite testing of two hypotheses, we have $\Theta(0)$ and $\Theta(1)$ that form a *partition* of the parameter space $\Theta$:

$$\Theta(0) \cup \Theta(1) = \Theta, \quad \Theta(0) \cap \Theta(1) = \emptyset$$

and that we wish to identify *which* of the two hypotheses is true:

$$\mathcal{H}_0 : \Theta \in \Theta(0) \quad \text{null hypothesis} \quad \text{versus} \quad \mathcal{H}_1 : \Theta \in \Theta(1) \quad \text{alternative hypothesis}.$$ 

Here, we adopt the classical Neyman-Pearson approach: given an upper bound $\alpha$ on the false-alarm probability, maximize the detection probability.

The fact that $\mathcal{H}_0$ is composite means that the false-alarm probability for a rule $\phi(X)$ is a function of $\theta$:

$$P_{FA}(\phi(X), \theta)$$

where $\theta \in \Theta(0)$. Therefore, to satisfy the upper bound $\alpha$, we
consider all tests $\phi(X)$ such that
\[ \max_{\theta \in \text{sp}_\Theta(0)} P_{FA}(\phi(X), \theta)) \leq \alpha. \] (18)

In this context,
\[ \max_{\theta \in \text{sp}_\Theta(0)} P_{FA}(\phi(X), \theta) \] (19)
is typically referred to as the size of the test $\phi(X)$. Therefore, the condition (18) states that we focus on tests whose size is upper-bounded by $\alpha$.

**Definition.** Among all tests $\phi(X)$ whose size is upper-bounded by $\alpha$ [i.e. (18) holds], we say that $\phi_{UMP}(X)$ is a uniformly most powerful (UMP) test if it satisfies
\[ P_D(\phi_{UMP}(X), \theta) \geq P_D(\phi(X), \theta) \]
for all $\theta \in \text{sp}_\Theta(1)$.

This is a very strong statement and very few hypothesis-testing problems have UMP tests. Note that Neyman-Pearson tests for simple hypotheses are UMP.

Hence, to find an UMP test for composite hypotheses, we need to first write a likelihood ratio for the simple hypothesis test with $\text{sp}_\Theta(0) = \{\theta_0\}$, $\text{sp}_\Theta(1) = \{\theta_1\}$, and $\text{sp}_\Theta = \{\theta_0, \theta_1\}$ and then transform this likelihood ratio in such a way that unknown quantities (e.g. $\theta_0$ and $\theta_1$) disappear from the test statistic.
(1) If such a transformation can be found, there is hope that a UMP test exists.

(2) However, we still need to figure out how to set a decision threshold ($\tau$, say) such that the upper bound (18) is satisfied.
Example 1: Detecting a Positive DC Level in AWGN (versus zero DC level)

Consider the following composite hypothesis-testing problem:

\[ H_0 : \quad \theta = 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(0) = \{0\} \quad \text{versus} \]
\[ H_1 : \quad \theta > 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(1) = (0, +\infty) \]

where the measurements \( X[0], X[1], \ldots, X[N - 1] \) are conditionally independent, identically distributed (i.i.d.) given \( \Theta = \theta \), modeled as

\[ \{X[n] \mid \Theta = \theta\} = \theta + W[n] \quad n = 0, 1, \ldots, N - 1 \]

with \( W[n] \) a zero-mean white Gaussian noise with known variance \( \sigma^2 \), i.e.

\[ W[n] \sim \mathcal{N}(0, \sigma^2) \]

implying

\[ f_{X \mid \Theta}(x \mid \theta) = \frac{1}{\sqrt{(2 \pi \sigma^2)^N}} \cdot \exp \left[ -\frac{1}{2 \sigma^2} \sum_{n=0}^{N-1} (x[n] - \theta)^2 \right] \quad (20) \]

where \( x = [x[0], x[1], \ldots, x[N - 1]]^T \). A sufficient statistic for \( \theta \) is

\[ \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x[n]. \]
Now, find the pdf of $\bar{x}$ given $\Theta = \theta$:

$$f_{\bar{X} | \Theta}(\bar{x} | \theta) = \mathcal{N}(\bar{x} | \theta, \sigma^2/N). \quad (21)$$

We start by writing the classical Neyman-Pearson test for the simple hypotheses with $\text{sp}_\Theta^{\text{simple}}(0) = \{0\}$ and $\text{sp}_\Theta^{\text{simple}}(1) = \{\theta_1\}, \theta_1 \in (0, +\infty)$:

$$\frac{f_{\bar{X} | \Theta}(\bar{x} | \theta_1)}{f_{\bar{X} | \Theta}(\bar{x} | 0)} = \frac{(2 \pi \sigma^2/N)^{-1/2} \cdot \exp[-\frac{1}{2 \sigma^2/N} (\bar{x} - \theta_1)^2]}{(2 \pi \sigma^2/N)^{-1/2} \cdot \exp[-\frac{1}{2 \sigma^2/N} (\bar{x})^2]} \geq \lambda.$$

Taking log etc. leads to

$$\theta_1 \bar{x} \geq \eta.$$

Since we know that $\theta_1 > 0$, we can divide both sides of the above expression by $\theta_1$ and accept $\mathcal{H}_1$ if

$$\phi(\bar{x}) : \bar{x} \geq \tau.$$

Hence, we transformed our likelihood ratio in such a way that $\theta_1$ disappears from the test statistic, i.e. we accomplished (1) above.

Now, on to (2). How to determine the threshold $\tau$ such that the upper bound (18) is satisfied? Based on (25), we know:

$$f_{\bar{X} | \Theta}(\bar{x} | 0) = \mathcal{N}(\bar{x} | 0, \sigma^2/N)$$
and, therefore,

\[
P_{FA}(\phi(X), 0) = \Pr_{X|\Theta} \{X > \tau \mid 0\}
\]

\[
= \Pr_{X|\Theta} \left\{ \frac{X - 0}{\sqrt{\sigma^2/N}} > \frac{\tau}{\sqrt{\sigma^2/N}} \mid 0 \right\}
\]

standard normal random var.

\[
= Q\left(\frac{\tau}{\sqrt{\sigma^2/N}}\right).
\]

Note that

\[
\max_{\theta \in \text{sp} \Theta(0)} P_{FA}(\phi(X), \theta) = P_{FA}(\phi(X), 0) = Q\left(\frac{\tau}{\sqrt{\sigma^2/N}}\right) = \alpha
\]

see (18) and (19). The most powerful test is achieved if the upper bound \(\alpha\) in (18) is reached by equality:

\[
\tau = \sqrt{\frac{\sigma^2}{N}} \cdot Q^{-1}(\alpha).
\]

Hence, we have accomplished (2), since this \(\tau\) yields exactly size \(\alpha\) for our test \(\phi(X)\).

To study the performance of the above test, we substitute
\[ (22) \] into the power function:

\[
\Pr_{X|\Theta}\{\bar{X} > \tau \mid \theta\} = \Pr_{X|\Theta}\left\{ \left. \frac{\bar{X} - \theta}{\sqrt{\sigma^2/N}} > \frac{\tau - \theta}{\sqrt{\sigma^2/N}} \right\} \right\}
\]

\[ = Q\left(\frac{\tau - \theta}{\sqrt{\sigma^2/N}}\right) = Q\left(\frac{1}{Q^{-1}(\alpha)} - \frac{\theta}{\sqrt{\sigma^2/N}}\right). \quad (23) \]
Consider the following composite hypothesis-testing problem:

\[ H_0 : \quad \theta \leq 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(0) = (-\infty, 0] \quad \text{versus} \]

\[ H_1 : \quad \theta > 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(1) = (0, +\infty) \]

where the measurements \( X[0], X[1], \ldots, X[N - 1] \) are conditionally i.i.d. given \( \Theta = \theta \), modeled as

\[ \{X[n] | \Theta = \theta\} = \theta + W[n] \quad n = 0, 1, \ldots, N - 1 \]

with \( W[n] \) a zero-mean white Gaussian noise with known variance \( \sigma^2 \), i.e.

\[ W[n] \sim \mathcal{N}(0, \sigma^2) \]

implying

\[
f_{X|\Theta}(x | \theta) = \frac{1}{\sqrt{(2 \pi \sigma^2)^N}} \cdot \exp \left[ -\frac{1}{2 \sigma^2} \sum_{n=0}^{N-1} (x[n] - \theta)^2 \right] \tag{24}\]

where \( x = [x[0], x[1], \ldots, x[N - 1]]^T \). A sufficient statistic for \( \theta \) is

\[
\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x[n].
\]
and
\[
f_{X|\Theta}(\bar{x}|\theta) = \mathcal{N}(\bar{x}|\theta, \sigma^2/N).
\] (25)
We start by writing the classical Neyman-Pearson test for the simple hypotheses with \(\Theta_{\text{simple}}(0) = \{\theta_0\}\) and \(\Theta_{\text{simple}}(1) = \{\theta_1\}\), where \(\theta_0 \in (-\infty, 0]\) and \(\theta_1 \in (0, +\infty)\), implying
\[
\frac{f_{X|\Theta}(\bar{x}|\theta_1)}{f_{X|\Theta}(\bar{x}|\theta_0)} = \frac{(2\pi \sigma^2/N)^{-1/2} \cdot \exp[-\frac{1}{2\sigma^2/N} (\bar{x} - \theta_1)^2]}{(2\pi \sigma^2/N)^{-1/2} \cdot \exp[-\frac{1}{2\sigma^2/N} (\bar{x} - \theta_0)^2]} \geq \lambda
\]
and
\[
\theta_0 < \theta_1.
\]
Taking log etc. leads to
\[
(\theta_1 - \theta_0) \bar{x} \geq \eta
\]
and, since \(\theta_0 < \theta_1\), to
\[
\phi(\bar{x}) : \bar{x} \geq \tau.
\]
Hence, we transformed our likelihood ratio in such a way that \(\theta_0\) and \(\theta_1\) disappear from the test statistic, i.e. we accomplished (1) above.

The power function of this test is
\[
Pr_{X|\Theta}\{\bar{X} > \tau | \theta\} = Pr_{X|\Theta}\left\{\frac{\bar{X} - \theta}{\sigma/\sqrt{N}} > \frac{\tau - \theta}{\sigma/\sqrt{N}} \bigg| \theta\right\} = Q\left(\frac{\tau - \theta}{\sigma/\sqrt{N}}\right)
\]
which is an increasing function of $\theta$. Recall the definition (19) of test size:

$$\max_{\theta \in \text{sp}_\Theta(0)} P_{\text{FA}}(\phi(X), \theta) = \max_{\theta \in \text{sp}_\Theta(0)} \Pr_X \{ \bar{X} > \tau | \theta \}$$

$$= \max_{\theta \in (-\infty, 0]} Q\left(\frac{\tau - \theta}{\sigma / \sqrt{N}}\right) = Q\left(\frac{\tau}{\sigma / \sqrt{N}}\right).$$

The most powerful test is achieved if the upper bound $\alpha$ in (18) is reached by equality:

$$\tau = \frac{\sigma}{\sqrt{N}} Q^{-1}(\alpha).$$

Hence, we have accomplished (2), since this $\tau$ yields exactly size $\alpha$ for our test $\phi(X)$.

![Power as a function of $\Theta$](image)

**FIGURE 10.1.** The power function for Example 2. The size of the test is the largest probability of rejecting $H_0$ when $H_0$ is true. This occurs at $\Theta = 0$, hence the size is $\text{power}(0)$. We choose the critical value $c$ so that $\text{power}(0) = \alpha$. 

EE 527, Detection and Estimation Theory, # 5c
Example 3: Detecting a Completely Unknown DC Level in AWGN

Consider now the composite hypothesis-testing problem:

\[ \mathcal{H}_0 : \quad \theta = 0 \quad \text{i.e.} \quad \theta \in \text{sp}_\Theta(0) = \{0\} \quad \text{versus} \]

\[ \mathcal{H}_1 : \quad \theta \neq 0 \quad \text{i.e.} \quad \theta \in \text{sp}_\Theta(1) = (-\infty, +\infty) \setminus \{0\} \]

where the measurements \( X[0], X[1], \ldots, X[N - 1] \) are conditionally i.i.d. given \( \Theta = \theta \), following

\[
 f_{X | \Theta}(x | \theta) = \frac{1}{\sqrt{(2\pi \sigma^2)^N}} \cdot \exp \left[ -\frac{1}{2 \sigma^2} \sum_{n=0}^{N-1} (x[n] - \theta)^2 \right]
\]

and \( x = [x[0], x[1], \ldots, x[N - 1]]^T \). A sufficient statistic for \( \theta \) is \( \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x[n] \) and the pdf of \( \bar{x} \) given \( \Theta = \theta \) is

\[
 f_{\bar{X} | \Theta}(\bar{x} | \theta) = \mathcal{N}(\bar{x} | \theta, \sigma^2 / N). \quad (26)
\]

We start by writing the classical Neyman-Pearson test for the simple hypotheses with \( \text{sp}_\Theta(0) = \{0\} \) and \( \text{sp}_\Theta(1) = \{\theta_1 \neq 0\} \):

\[
 \theta_1 \bar{x} > \eta.
\]

We cannot accomplish (1), since \( \theta_1 \) cannot be removed from the test statistic; therefore, UMP test does not exist for the above problem.
Consider a scalar parameter $\theta$. We say that $f_{X \mid \Theta}(x \mid \theta)$ belongs to the monotone likelihood ratio (MLR) family if the pdfs (or pmfs) from this family

- satisfy the identifiability condition for $\theta$ (i.e. these pdfs are distinct for different values of $\theta$) and
- there is a scalar statistic $T(x)$ such that, for $\theta_0 < \theta_1$, the likelihood ratio

$$\Lambda(x; \theta_0, \theta_1) = \frac{f_{X \mid \Theta}(x \mid \theta_1)}{f_{X \mid \Theta}(x \mid \theta_0)}$$

is a monotonically increasing function of $T(x)$.

If $f_{X \mid \Theta}(x \mid \theta)$ belongs to the MLR family, then use the following test:

$$\phi_\lambda(x) = \begin{cases} 
1, & \text{for } T(x) \geq \lambda, \\
0, & \text{for } T(x) < \lambda
\end{cases}$$
and set

$$\alpha = P_{FA}(\phi(X), \theta_0) = \Pr_{X|\Theta}\{T(X) \geq \lambda \mid \theta_0\} \quad (27)$$

e.g. use this condition to find the threshold $\lambda$.

This test has the following properties:

**(i)** With $\alpha$ given by (27), $\phi_{\lambda}(x)$ is UMP test of size $\alpha$ for testing

$$\mathcal{H}_0 : \theta > \theta_0 \quad \text{versus} \quad \mathcal{H}_1 : \theta \leq \theta_0.$$

**(ii)** For each $\lambda$, the power function

$$\Pr_{X|\Theta}\{T(X) \geq \lambda \mid \theta\} \quad (28)$$

is a monotonically increasing function of $\theta$.

**Note:** Consider the one-parameter exponential family

$$f_{X|\Theta}(x \mid \theta) = h(x) \exp[\eta(\theta) T(x) - B(\theta)]. \quad (29)$$

Then, if $\eta(\theta)$ is a monotonically increasing function of $\theta$, the class of pdfs (pmfs) (29) satisfies the MLR conditions.
Example: Detection for Exponential Random Variables

Consider conditionally i.i.d. measurements $X[0], X[1], \ldots, X[N-1]$ given the parameter $\theta > 0$, following the exponential pdf:

$$f_{X|\Theta}(x[n] | \theta) = \text{Expon}(x[n] | 1/\theta) = \frac{1}{\theta} \exp(-\theta^{-1} x[n]) i_{(0, +\infty)}(x[n]).$$

The likelihood function of $\theta$ for all observations $x = [x[0], x[1], \ldots, x[N-1]]^T$ is

$$f_{X|\Theta}(x | \theta) = \frac{1}{\theta^N} \exp[-\theta^{-1} T(x)] \prod_{n=0}^{N-1} i_{(0, +\infty)}(x[n])$$

where

$$T(x) = \sum_{n=0}^{N-1} x[n].$$

Since $f_{X|\Theta}(x | \theta)$ belongs to the one-parameter exponential family (29) and $\eta(\theta) = -\theta^{-1}$ is a monotonically increasing function of $\theta$. Therefore, the test

$$\phi_\lambda(x) = \begin{cases} 1, & \text{for } T(x) \geq \lambda, \\ 0, & \text{for } T(x) < \lambda \end{cases}$$
is UMP for testing

\[ \mathcal{H}_0 : \theta > \theta_0 \quad \text{versus} \quad \mathcal{H}_1 : \theta \leq \theta_0. \]

The sum of i.i.d. exponential random variables follows the Erlang pdf (which is a special case of the gamma pdf):

\[
f_{T \mid \Theta}(T \mid \theta) = \frac{1}{\theta^N} \frac{T^{N-1}}{(N-1)!} \exp\left(-\frac{T}{\theta}\right) i_{(0, +\infty)}(T) = \text{Gamma}(T \mid N, \theta^{-1}).
\]

Therefore, the size of the test can be written as

\[
\alpha = \Pr_{X \mid \Theta} \{ T(X) \geq \lambda \mid \theta_0 \} = \frac{1}{\theta_0^N} \int_{\lambda}^{+\infty} \frac{t^{N-1}}{(N-1)!} \exp\left(-\frac{t}{\theta_0}\right) dt
\]

\[
= \left[ 1 + \frac{\lambda}{\theta_0} + \cdots + \frac{1}{(N-1)!} \left(\frac{\lambda}{\theta_0}\right)^{N-1} \right] \exp\left(-\frac{\lambda}{\theta_0}\right)
\]

where the integral is evaluated using integration by parts. For \( N = 1 \), we have

\[ \lambda = \theta_0 \ln(1/\alpha). \]
Generalized Likelihood Ratio (GLR) Test

Recall again that, in composite testing of two hypotheses, we have $\Theta(0)$ and $\Theta(1)$ that form a partition of the parameter space $\Theta$:

$$\Theta(0) \cup \Theta(1) = \Theta, \quad \Theta(0) \cap \Theta(1) = \emptyset$$

and that we wish to identify which of the two hypotheses is true:

$$\mathcal{H}_0 : \theta \in \Theta(0) \quad \text{null hypothesis} \quad \text{versus} \quad \mathcal{H}_1 : \theta \in \Theta(1) \quad \text{alternative hypothesis.}$$

In GLR tests, we replace the unknown parameters by their maximum-likelihood (ML) estimates under the two hypotheses. Hence, accept $\mathcal{H}_1$ if

$$\Lambda_{GLR}(x) = \frac{\max_{\theta \in \Theta(1)} f(x | \Theta(0) \theta)}{\max_{\theta \in \Theta(0)} f(x | \Theta(0) \theta)} > \tau.$$ 

This test has no UMP optimality properties, but often works well in practice.
Example: Detecting a Completely Unknown DC Level in AWGN

Consider again the composite hypothesis-testing problem from p. 38:

\[ \mathcal{H}_0 : \quad \theta = 0 \quad \text{i.e.} \quad \theta \in \text{sp}_\Theta(0) = \{0\} \quad \text{versus} \]

\[ \mathcal{H}_1 : \quad \theta \neq 0 \quad \text{i.e.} \quad \theta \in \text{sp}_\Theta(1) = (-\infty, +\infty) \backslash \{0\} \]

where the measurements \( X[0], X[1], \ldots, X[N - 1] \) are conditionally i.i.d. given \( \Theta = \theta \), following

\[
f_{X|\Theta}(x|\theta) = \frac{1}{\sqrt{(2\pi \sigma^2)^N}} \cdot \exp \left[ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - \theta)^2 \right]
\]

and \( x = [x[0], x[1], \ldots, x[N - 1]]^T \). A sufficient statistic for \( \theta \) is \( \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x[n] \) and the pdf of \( \bar{x} \) given \( \Theta = \theta \) is

\[
f_{\bar{X}|\Theta}(\bar{x}|\theta) = \mathcal{N}(\bar{x}|\theta, \sigma^2/N).
\]

Our GLR test accepts \( \mathcal{H}_1 \) if

\[
\Lambda_{\text{GLR}}(x) = \frac{\max_{\theta \in \text{sp}_\Theta(1)} f_{X|\Theta}(x|\theta)}{f_{\bar{X}|\Theta}(\bar{x}|0)} > \tau.
\]

Now,

\[
\bar{x} = \arg \max_{\theta \in \text{sp}_\Theta(1)} f_{X|\Theta}(x|\theta)
\]
and

\[ f_{\bar{X} | \Theta}(\bar{x} | 0) = \mathcal{N}(\bar{x} | 0, \sigma^2/N) \]
\[ = \frac{1}{\sqrt{2 \pi \sigma^2/N}} \exp \left( -\frac{1}{2} \frac{\bar{x}^2}{\sigma^2/N} \right) \]
\[ f_{\bar{X} | \Theta}(\bar{x} | \bar{x}) = \mathcal{N}(\bar{x} | 0, \sigma^2/N) = \frac{1}{\sqrt{2 \pi \sigma^2/N}} \]

yielding

\[ \ln \Lambda_{GLR}(\bar{x}) = \frac{N \bar{x}^2}{2 \sigma^2}. \]

Therefore, we accept \( \mathcal{H}_1 \) if

\[ (\bar{x})^2 > \gamma \]

or

\[ |\bar{x}| > \eta. \]

We compare this detector with the (not realizable, also called clairvoyant) UMP detector that assumes the knowledge of the sign of \( \theta \) under \( \mathcal{H}_1 \). Assuming that the sign of \( \theta \) under \( \mathcal{H}_1 \) is known, we can construct the UMP detector, whose ROC curve is given by

\[ P_D = Q(Q^{-1}(P_{FA}) - d) \]

where \( d = \sqrt{N \theta^2 / \sigma^2} \) and \( \theta \) is the value of the parameter under \( \mathcal{H}_1 \); see (23) for the case where \( \theta > 0 \) under \( \mathcal{H}_1 \). All other detectors have \( P_D \) below this upper bound.
GLR test: Decide $\mathcal{H}_1$ if $|\bar{x}| > \eta$. To make sure that the GLR test is implementable, we must be able to specify a threshold $\eta$ so that the false-alarm probability is upper-bounded by a given size $\alpha$. This is possible in our example:

\[
P_{FA}(\phi(x), 0) = \Pr_{X|\Theta}\{|X| > \eta | 0\} \quad \text{see (26)}
\]
\[
\text{symmetry} \quad = \quad 2 \Pr_{X|\Theta}\{X > \eta | 0\} = 2 Q(\eta/\sqrt{\sigma^2/N})
\]
\[
P_D(\phi(x), \theta) = \Pr_{X|\Theta}\{|X| > \eta | \theta\}
\]
\[
= \quad \Pr_{X|\Theta}\{X > \eta | \theta\} + \Pr_{X|\Theta}\{X < -\eta | \theta\}
\]
\[
= \quad Q\left(\frac{\eta - \theta}{\sqrt{\sigma^2/N}}\right) + Q\left(\frac{\eta + \theta}{\sqrt{\sigma^2/N}}\right)
\]
\[
= \quad Q\left(Q^{-1}(\alpha/2) - \frac{\theta}{\sqrt{\sigma^2/N}}\right)
\]
\[
+ Q\left(Q^{-1}(\alpha/2) + \frac{\theta}{\sqrt{\sigma^2/N}}\right).
\]

In this case, GLR test is only slightly worse than the clairvoyant detector (Figure 6.4 in Kay-II):
Example: DC level in WGN with $A$ and $\sigma^2$ both unknown. Recall that $\sigma^2$ is called a nuisance parameter since we care exclusively about $\theta$. Here, the GLR test for

$\mathcal{H}_0 : \theta = 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(0) = \{0\}$ versus $\mathcal{H}_1 : \theta \neq 0 \quad \text{i.e. } \theta \in \text{sp}_\Theta(1) = (-\infty, +\infty) \setminus \{0\}$

accepts $\mathcal{H}_1$ if

$$\Lambda_{\text{GLR}}(x) = \frac{\max_{\theta, \sigma^2} f_X | \Theta, \Sigma^2(x | \theta, \sigma^2)}{\max_{\sigma^2} f_X | \Theta, \Sigma^2(x | 0, \sigma^2)} > \gamma$$
where

\[ f_{\mathbf{x} | \Theta, \Sigma^2}(\mathbf{x} | \theta, \sigma^2) = \frac{1}{\sqrt{(2 \pi \sigma^2)^N}} \cdot \exp \left[ - \frac{1}{2 \sigma^2} \sum_{n=0}^{N-1} (x[n] - \theta)^2 \right]. \]  

(30)

Here,

\[
\max_{\theta, \sigma^2} f_{\mathbf{x} | \Theta, \Sigma^2}(\mathbf{x} | \theta, \sigma^2) = \frac{1}{[2 \pi \hat{\sigma}^2_0(\mathbf{x})]^{N/2}} \cdot e^{-N/2}
\]

\[
\max_{\sigma^2} f_{\mathbf{x} | \Theta, \Sigma^2}(\mathbf{x} | 0, \sigma^2) = \frac{1}{[2 \pi \hat{\sigma}^2_0(\mathbf{x})]^{N/2}} \cdot e^{-N/2}
\]

where

\[
\hat{\sigma}^2_0(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} x^2[n]
\]

\[
\hat{\sigma}^2_1(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} (x[n] - \overline{x})^2.
\]

Hence,

\[
\Lambda_{\text{GLR}}(\mathbf{x}) = \left( \frac{\hat{\sigma}^2_0(\mathbf{x})}{\hat{\sigma}^2_1(\mathbf{x})} \right)^{N/2}
\]

i.e. GLR test fits data with the “best” DC-level signal \( \hat{\theta}_{\text{ML}} = \overline{x} \), finds the residual variance estimate \( \hat{\sigma}^2_1 \), and compares this estimate with the variance estimate \( \hat{\sigma}^2_0 \) under the null case (i.e.
for \( \theta = 0 \). When sufficiently strong signal is present, \( \hat{\sigma}_1^2 \ll \hat{\sigma}_0^2 \) and \( \Lambda_{\text{GLR}}(x) \gg 1 \).

Note that

\[
\hat{\sigma}_1^2(x) = \frac{1}{N} \sum_{n=1}^{N} (\bar{x} - x[n])^2
\]

\[
= \frac{1}{N} \sum_{n=1}^{N} (x^2[n] - 2 \bar{x} x[n] + \bar{x}^2)
\]

\[
= \left( \frac{1}{N} \sum_{n=1}^{N} x^2[n] \right) - 2 \bar{x}^2 + \bar{x}^2
\]

\[
= \hat{\sigma}_0^2(x) - \bar{x}^2.
\]

Hence,

\[
2 \ln \Lambda_{\text{GLR}}(x) = N \ln \left( \frac{\hat{\sigma}_0^2(x)}{\hat{\sigma}_0^2(x) - \bar{x}^2} \right) = N \ln \left( \frac{1}{1 - \bar{x}^2/\hat{\sigma}_0^2(x)} \right).
\]

Note that

\[
0 \leq \frac{\bar{x}^2}{\hat{\sigma}_0^2(x)} \leq 1
\]

and \( \ln[1/(1 - z)] \) is monotonically increasing on \( z \in (0, 1) \). Therefore, an equivalent test can be constructed as follows:

\[
T(x) = \frac{\bar{x}^2}{\hat{\sigma}_0^2(x)} > \tau.
\]
The pdf of $T(X)$ given $\theta = 0$ does not depend on $\sigma^2$ and, therefore, GLR test can be implemented, i.e. it is CFAR.

**Definition.** A test is constant false alarm rate (CFAR) if we can find a threshold that yields a test whose size is equal to $\alpha$.

In other words, we should be able to set the threshold independently of the unknown parameters, i.e. the distribution of the test statistic under $\mathcal{H}_0$ does not depend on the unknown parameters.