Minimum probability of error or MAP decision rule is given by:
\[ P_{m} \Lambda_{m}(\vec{R}) \gtrless n_{m}^{\text{not}} P_{k} \Lambda_{k}(\vec{R}) \]
- decision: \( \max_{[k]} P_{k} \Lambda_{k}(\vec{R}) \).
If a priori probabilities i.e. \( P_{i} \)'s are not known, one option would be to assume equally probable hypotheses; i.e. \( P_{i} = \frac{1}{M}, i = 0, 1, \cdots, M - 1 \).
For this choice of a priori probabilities, the test becomes
\[ \Lambda_{m}(\vec{R}) \gtrless n_{m}^{\text{not}} \Lambda_{k}(\vec{R}) \]
or
\[ p_{M}(\vec{R}) \gtrless n_{m}^{\text{not}} p_{K}(\vec{R}) \]
- decision: \( \max_{[k]} p_{k}(\vec{R}) \) i.e. choose the hypothesis that yields the largest \( p_{k}(\vec{R}) \) or \( \Lambda_{k}(\vec{R}) \). This receiver is known as the maximum likelihood (ML) receiver or detector.
We showed that the decision threshold in the binary hypothesis testing, i.e.
\[ \Lambda(\vec{R}) \gtrless n_{0}^{\text{select}} \frac{P_{0}(C_{10} - C_{00})}{P_{1}(C_{01} - C_{11})} \triangleq \eta \]
depends on the a priori probability \( P_{0} \).
Thus the statistic used for the decision making, i.e. \( \Lambda(\vec{R}) \) is invariant in a priori probability \( P_{0} \). This implies that the sufficient statistic is unaffected for the ML detector. Only the threshold \( \eta \) varies with \( P_{0} \).

Minimax detector
This is also used for the case of unknown a priori probabilities.
For a given channel, the risk is a function of \( P_{i} \forall i = 0, 1, \cdots, M - 1 \).
We can inspect all possible risk functions \( R(\vec{P}) \) in the domain of \( \vec{P} \). We then choose the \( \vec{P} \) values that maximize the risk function (worst case scenario), and set up the decision based on that a priori vector.

- decision: \( \min[\max_{\vec{P}} R(\vec{P})] = \min[\Re(\vec{P}_{\text{max}})] \).
- example: Binary decision

\[ \Lambda(\vec{R}) \gtrless n_{0}^{\text{select}} \eta = \frac{P_{0}(C_{10} - C_{00})}{P_{1}(C_{01} - C_{11})} \]

For minimax detector, use

\[ \Lambda(\vec{R}) \gtrless n_{0}^{\text{select}} \eta^{*} = \frac{P_{0}^{*}(C_{10} - C_{00})}{P_{1}^{*}(C_{01} - C_{11})} \]

where \( P_{0}^{*} = (1 - P_{0}^{*}) \)

Consider:
\[ H_{0} : r_{i} \sim \text{N}(0, \sigma_{0}^{2}) \]
\[ H_{1} : r_{i} \sim \text{N}(0, \sigma_{1}^{2}) \]
where \( i = 1, \cdots, N \) and \( \sigma_1 > \sigma_0 \). Observation pdf under \( H_K, K = 0, 1 \) is:

\[
p(R|H_K) = \frac{1}{(\sqrt{2\pi}\sigma_K)^N} e^{\exp} \left[- \frac{\sum_{i=1}^{N} r_i^2}{2\sigma_K^2} \right]
\]

Likelihood function:

\[
\lambda(R) = \frac{p(R|H_1)}{p(R|H_0)} = \left( \frac{\sigma_0}{\sigma_1} \right)^N e^{\exp} \left[ \frac{1}{2} \left( \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} \right) \sum_{i=1}^{N} r_i^2 \right] \geq_{select H_0} \eta
\]

Log-likelihood function:

\[
\ln \lambda(R) = N \ln \left( \frac{\sigma_0}{\sigma_1} \right) + \left[ \frac{1}{2} \left( \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} \right) \sum_{i=1}^{N} r_i^2 \right] \geq_{select H_0} \ln \eta
\]

Thus the sufficient statistic can be expressed via:

\[
l(R) = \frac{1}{\overline{R}^2} \sum_{i=1}^{N} r_i^2 \geq_{select H_0} \gamma \ln \eta + N \ln \frac{\sigma_0}{\sigma_1} \frac{1}{2} \sigma_0^2 \left( \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \left( \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} \right) \geq_{select H_0} \gamma
\]

Test: General Bayes decision

\[
\overline{R}^2 \geq_{select H_1} \gamma
\]

The r.v. \( l(R) = \overline{R}^2 \) is central chi-square distributed with \( N \) degrees of freedom. The distribution is central since \( r_i \)'s have zero mean.

\( E(R^2|H_K) = \sigma_K^2 \), for \( K = 0, 1 \). Variance of \( \overline{R}^2 \) decreases as \( N \) increases.

**Minimum probability of error as a function of the a priori probabilities**

In general, probability of error \( P_e \) is given as:

\[
P_e = P_0P_F + P_1P_M
\]

\[
= P_0P_F + (1 - P_0)P_M
\]

\[
= P_M + P_0(P_F - P_M)
\]

This expression indicates that the probability of error \( P_e \) is a linear function of \( P_0 \) if both \( P_F \) and \( P_M \) are invariant in \( P_0 \).

Note that in a Bayes decision rule, the threshold and as a result \( P_F \) and \( P_M \) are functions of \( P_0 \) and nonlinear as shown in figure 1.

In a general Bayes rule problem, \( P_F \) and \( P_M \) are defined as:

\[
P_F = \int_{-\infty}^{\gamma(P_0)} p(l|H_0)dl
\]

\[
P_M = \int_{-\infty}^{\gamma(P_0)} p(l|H_1)dl
\]

Minimum probability of error as a function of a priori probabilities is given by:

\[
P_e = P_0P[l > \gamma(P_0)|H_0] + (1 - P_0)P[l < \gamma(P_0)|H_1]
\]

where \( \gamma(P_0) \triangleq g[\eta(P_0)] = g(\frac{P_0}{1-P_0}) \).

\( \eta(P_0) = \frac{P_0}{1-P_0} \) with \( C_{00} = C_{11} = 0 \) and \( C_{10} = C_{01} = 1 \).
Suppose $P_{e_{\text{min}}}$ takes on its maximum value at $P_0 = P_0^\ast$ (i.e. minimax detector with unitary cost) as shown in figure (2).
Select the the threshold based on $P_0 = P_0^\ast$.
Therefore, $\eta(P_0^\ast) = \frac{P_0^\ast}{1-P_0^\ast}$ and $\gamma(P_0^\ast) = g(\frac{P_0^\ast}{1-P_0^\ast})$. Once a threshold is selected based on $P_0 = P_0^\ast$, then the resultant probability of error for the actual $P_0$ is:

$$P_e^\ast(P_0) = P_0^\ast P_F^\ast + (1-P_0^\ast)P_M^\ast$$

$$= P_M^\ast + P_0^\ast(P_F^\ast - P_M^\ast)$$

which is a linear function of $P_0$. 

**Figure 1.** Plot of $P_0$ vs. $P_e$ for Bayes decision rule

**Figure 2.** Plot of $P_0$ vs. $P_{e_{\text{min}}}$
Consider the curve for $P_e^*$ to be as shown in figure 3. We know that $P_{e_{\text{min}}}(P_0)$ should always yield the minimum probability of error. But from figure 3, we have $P_{e_{\text{min}}}^*(\alpha) > P_e^*(\alpha)$ which is not possible. Thus the distribution of the line $P_e^*(P_0)$ should be above the curve $P_{e_{\text{min}}}(P_0)$ at every point except one where the two curves touch each other. In other words, the line for $P_e^*(P_0)$ is a tangent to the latter curve as shown in figure 4.
This implies that $P^*_e(P_0) = \text{constant}$, i.e. not varying with $P_0$

Therefore

$$P^*_e(P_0) = P^*_M + P_0(P^*_F - P^*_M)$$

$$\Rightarrow P^*_F = P^*_M$$

at the minimax point.

**Neyman Pearson criterion**

We wish to construct the decision problem based on the detection and false alarm probabilities.

- **Criterion**: Fix $P_F = \alpha$ and maximize $P_D$.

- **Solution**: Define the Lagrange:

  $$\mathcal{L} = P_D - \lambda(P_F - \alpha)$$

  where $\lambda$ is the Lagrange multiplier. - Substitute for $P_D$ and $P_F$ based on the pdf’s and the decision threshold.

  $$\mathcal{L} = \int_{Z_1} p(\vec{R}|H_1) d\vec{R} - \lambda \int_{Z_1} p(\vec{R}|H_0) d\vec{R} - \alpha$$

  $$= \lambda \alpha + \int_{Z_1} [p(\vec{R}|H_1) - \lambda(p(\vec{R}|H_0))] d\vec{R}$$

**Case 1**: $\lambda < 0$

In this case, the integrand $[p(\vec{R}|H_1) - \lambda(p(\vec{R}|H_0))]$ is always positive. Thus the function $\mathcal{L}$ is maximized when the integration is done over the largest possible region for $Z$, i.e $Z_1 = Z$.

$\Rightarrow$ choose $H_1$ for all $\vec{R}$, which is not an acceptable solution.

**Case 2**: $\lambda > 0$

To achieve maximum $\mathcal{L}$, we should integrate over the region in $Z$ where the integrand is positive. i.e. $[p(\vec{R}|H_1) - \lambda(p(\vec{R}|H_0))] > 0$ for $\vec{R} \in Z_1$.

This yields the following decision rule:

$$p(\vec{R}|H_1) \gtrless \begin{cases} 1 & \text{select } H_1 \\ 0 & \text{select } H_0 \end{cases} \lambda p(\vec{R}|H_0)$$

or $\Lambda(\vec{R}) = \begin{cases} 1 & \text{select } H_1 \\ 0 & \text{select } H_0 \end{cases} \lambda$

This is the Bayes decision rule with

$$\lambda \triangleq \eta = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}$$

To find $\lambda$ (the threshold), we use the fact that $P_F = \alpha$, implying

$$P_F = \int_{\lambda}^{\infty} p(\Lambda|H_0) d\lambda$$

Therefore, we vary $\lambda$ until $P_F = \alpha$ is achieved.

**Receiver Operating Characteristics (ROC)**

ROC is simply a plot of $P_D$ vs $P_F$ for a given receiver (decision rule) as a function of the parameter of interest. Likelihood ratio test (LRT):

$$\lambda = \begin{cases} 1 & \text{select } H_1 \\ 0 & \text{select } H_0 \end{cases} \eta$$

yields $P_F$ and $P_D$ as functions of parameters such as additive noise variance or average SNR, signaling type (unipolar, bipolar) etc.
Example

\[ H_0 : \ r_i = A_0 + n_i \quad i = 1, 2, ..., N \]

\[ H_1 : \ r_i = A_1 + n_i \]

where \((A_0, A_1)\) are constants and \(n_i \sim (0, \sigma_n^2)\).

We showed that

\[
\Lambda(\tilde{R}) = \frac{p(\tilde{R}|H_1)}{p(\tilde{R}|H_0)} = \exp\left[\frac{\sum_{i=1}^N (r_i - A_0)^2 - (r_i - A_1)^2}{2\sigma_n^2}\right]
\]

For general ASK signaling, we have:

\[
l(\tilde{R}) = \sum_{i=1}^N (S_1 - S_0) r_i \underbrace{\sim}_{\text{select } H_i} \ln \eta \cdot \sigma_n^2 + \frac{1}{2} \sum_{i=1}^N (S_{1i}^2 - S_{0i}^2)
\]

For this example, we have \(S_{0i} = A_0\) and \(S_{1i} = A_1\), \(\forall i = 1, 2, ..., N\). Thus the test becomes:

\[
l(\tilde{R}) = \sum_{i=1}^N (A_1 - A_0) r_i \underbrace{\sim}_{\text{select } H_0} \ln \eta \cdot \sigma_n^2 + \frac{N}{2} (A_1^2 - A_0^2)
\]

Redefine the sufficient statistic by normalizing it as:

\[
l(\tilde{R}) \triangleq \sum_{i=1}^N r_i \sqrt{\frac{\ln \eta \cdot \sigma_n}{\sqrt{N}(A_1 - A_0)}} + \frac{\sqrt{N}(A_1 + A_0)}{2\sigma_n}
\]

We have the following distributions which are also shown in figure 5:

\[
l \mid H_0 \sim N\left(\frac{\sqrt{N}A_0}{\sigma_n}, 1\right)
\]

\[
l \mid H_1 \sim N\left(\frac{\sqrt{N}A_1}{\sigma_n}, 1\right)
\]

![Figure 5](image-url)
Define \( d = \frac{\sqrt{N}(A_1 - A_0)}{\sigma_n} \) and \( D = \frac{\sqrt{N}(A_1 + A_0)}{\sigma_n} \). Therefore, the LRT becomes:

\[
l(\vec{R}) \triangleq \sum_{i=1}^{N} \frac{r_i}{\sqrt{N}\sigma_n} \gtrsim \text{select}\ n_1 \ln \frac{\eta}{d} + \frac{D}{2}
\]

Note: For unipolar ASK where \( A_0 = 0 \) and \( D = d \), the quantity representing average bit energy under \( P_0 = P_1 = \frac{1}{2} \) becomes:

\[
E_b = \frac{1}{2} (A - d^2) = \frac{\sigma_n^2}{4N}(D_2 + d^2)
\]

Thus, with \((d, D)\), we identify two important features of this form of data transmission.

a) Separation of the two hypotheses in the \( l \) domain;

b) Average energy of the transmitter with respect to the noise power.

Performance probabilities:

\[
P_F = \int_{\frac{\ln n}{d} + \frac{D}{2} - \frac{\sqrt{N}A_0}{\sigma_n}}^{\infty} p(l|H_0)dl
\]

\[
= \int_{\frac{\ln n}{d} + \frac{D}{2} - \frac{\sqrt{N}A_0}{\sigma_n}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(l - \frac{\sqrt{N}A_0}{\sigma_n})^2}{2}} dl
\]

\[
= \text{erfc}\left(\frac{\ln n}{d} + \frac{D}{2} - \frac{\sqrt{N}A_0}{\sigma_n}\right)
\]

Similarly

\[
P_D = \int_{\frac{\ln n}{d} + \frac{D}{2} - \frac{\sqrt{N}A_0}{\sigma_n}}^{\infty} p(l|H_1)dl
\]

\[
= \text{erfc}\left(\frac{\ln n}{d} + \frac{D}{2} - \frac{\sqrt{N}A_1}{\sigma_n}\right)
\]

Property: The threshold \( \eta \) for the likelihood ratio test is given by:

\[
\eta = \left. \frac{dp_D}{dP_F} \right|_{\text{operating point of interest}}
\]

Proof: \( P_D = \int_{\eta}^{\infty} p(\Lambda|H_1)d\Lambda \)

\[
= \int_{Z_1} p(R|H_1)dR
\]

But, we know:

\[
\Lambda(R) = \frac{p(R|H_1)}{p(R|H_0)}
\]

or

\[
p(R|H_1) = \Lambda(R).p(R|H_0)
\]

We substitute this in the expression for \( P_D \):

\[
P_D = \int_{Z_1} \Lambda(R).p(R|H_0)dR
\]

Rewrite the above in terms of \( \Lambda|H_0 \):

\[
P_D = \int_{\eta}^{\infty} \Lambda(p|H_0)d\Lambda
\]
We also have:

\[ P_F = \int_\eta^{\infty} p(\Lambda|H_0) d\Lambda \]

Therefore, using Leibnitz rule of differentiation (refer to Papoulis & Pillai, page: 181)

\[ \frac{dP_D}{d\eta} = -\eta p(\eta|H_0) \]
\[ \frac{dP_F}{d\eta} = -p(\eta|H_0) \]

Finally:

\[ \frac{dP_D/d\eta}{dP_F/d\eta} = \frac{-\eta p(\eta|H_0)}{-p(\eta|H_0)} = \eta \]

\[ \Rightarrow \frac{dP_D}{dP_F} = \eta \]