<Ex>

H_1: receive $\underline{r}(\mu)$ with $r_i(\mu)$'s iid and $f(r_i|H_1) \sim G(0,\sigma_1^2)$

H_0: receive $\underline{r}(\mu)$ with $r_i(\mu)$'s iid and $f(r_i|H_0) \sim G(0,\sigma_0^2)$, $\sigma_1 > \sigma_0$

$$\Lambda(\underline{r}) = \frac{f(\underline{r}(\mu)|H_1)}{f(\underline{r}(\mu)|H_0)} = \frac{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{\frac{-r_i^2}{2\sigma_1^2}}}{\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{\frac{-r_i^2}{2\sigma_0^2}}}$$

$$\Leftrightarrow \ln\Lambda(\underline{r}) = \frac{1}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\sum_{i=1}^{N} r_i^2 + N\ln\frac{\sigma_0}{\sigma_1} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta_{Bayes}$$

$$\Leftrightarrow \ell(\underline{r}) \equiv \text{sufficient statistic} = \sum_{i=1}^{N} r_i^2$$

, and the equivalent test is

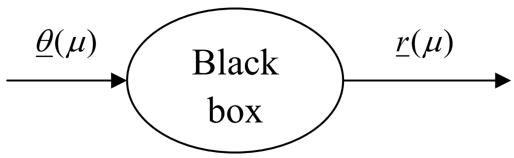
$$\ell(\underline{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \alpha \text{ with } \alpha \equiv \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2}\left(\ln\eta_{Bayes} - N\ln\frac{\sigma_0}{\sigma_1}\right)$$

Note: The $\ell(\underline{r}(\mu))$ here is a nonlinear function of $\{r_i(\mu)\}_{i=1}^{N}$

- More on Sufficient Statistics

  — A function of data is referred to as a statistic.

  — Let us consider



The observer aims to understand $\underline{\theta}(\mu)$ based on the observed $\underline{\mathrm{r}}(\mu)$.

— A statistic $\ell(\underline{\mathrm{r}}(\mu))$ is said to be sufficient if the *a posteriori* density of $\underline{\theta}(\mu)$, given $\ell(\underline{\mathrm{r}}(\mu))$, i.e., $f(\underline{\theta} \mid \ell(\underline{\mathrm{r}}))$, is equivalent to the *a posteriori* density of $\underline{\theta}(\mu)$, given $\underline{\mathrm{r}}(\mu)$, i.e., $f(\underline{\theta} \mid \underline{\mathrm{r}})$. That is,

$$f(\underline{\theta} \mid \ell(\underline{\mathrm{r}})) = f(\underline{\theta} \mid \underline{\mathrm{r}}), \ \forall\, \underline{\mathrm{r}}, \underline{\theta}$$

— Or, equivalently, a statistic $\ell(\underline{\mathrm{r}}(\mu))$ is sufficient if

$$f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}), \underline{\theta}) = f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}})), \ \forall\, \underline{\mathrm{r}}, \underline{\theta}$$

(That is, given a statistic $\ell(\underline{\mathrm{r}}(\mu))$, $\underline{\mathrm{r}}(\mu)$ does not provide any more information about $\underline{\theta}(\mu)$.)

<pf>

$$f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}), \underline{\theta}) = \frac{f(\underline{\mathrm{r}}, \ell(\underline{\mathrm{r}}), \underline{\theta})}{f(\ell(\underline{\mathrm{r}}), \underline{\theta})}$$

$$= \frac{f(\underline{\theta} \mid \underline{\mathrm{r}}, \ell(\underline{\mathrm{r}})) f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}})) f(\ell(\underline{\mathrm{r}}))}{f(\underline{\theta} \mid \ell(\underline{\mathrm{r}})) f(\ell(\underline{\mathrm{r}}))}$$

$$= \frac{f(\underline{\theta} \mid \underline{\mathrm{r}}, \ell(\underline{\mathrm{r}})) f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}))}{f(\underline{\theta} \mid \ell(\underline{\mathrm{r}}))} \qquad (+)$$

Since $\ell(\underline{\mathrm{r}})$ is a function of $\underline{\mathrm{r}}$,

$$f(\underline{\theta} \mid \underline{\mathrm{r}}, \ell(\mathrm{r})) = f(\underline{\theta} \mid \underline{\mathrm{r}}), \ \forall\, \underline{\mathrm{r}}, \underline{\theta}$$

It follows from (+) that

$$f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}), \underline{\theta}) = \frac{f(\underline{\theta} \mid \underline{\mathrm{r}}) f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}))}{f(\underline{\theta} \mid \ell(\underline{\mathrm{r}}))}$$

Now, if $f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}), \theta) = f(\underline{\mathrm{r}} \mid \ell(\underline{\mathrm{r}}))$, then

$(+) = f(\underline{r} \mid \ell(\mathrm{r})) \Rightarrow f(\underline{\theta} \mid \underline{\mathrm{r}}) = f(\underline{\theta} \mid \ell(\underline{\mathrm{r}})) \Rightarrow \ell(\underline{\mathrm{r}})$ is a sufficient statistic.

Q.E.D.

—Neyman-Fisher Factorization Theorem

A statistic $\ell(\underline{r}(\mu))$ is sufficient iff

$$f(\underline{r} \mid \underline{\theta}) = g(\underline{\theta}, \ell(\underline{r}))h(\underline{r}), \forall \underline{r}, \underline{\theta}$$

where $g(\underline{\theta}, \ell(\underline{r}))$ is a function of $\underline{\theta}$ and $\ell(\underline{r})$, and $h(\underline{r})$ is a function of $\underline{r}$ only.

<pf> For a rigorous proof, see S. Zacks, “The Theory of Statistical Inference,” Wiley, 1971.

"$\Leftarrow$" If $f(\underline{r}|\underline{\theta}) = g(\underline{\theta},\ell(\underline{r}))h(\underline{r})$, then

$$f(\underline{\theta}|\underline{r}) = \frac{f(\underline{r}|\underline{\theta})f(\underline{\theta})}{f(\underline{r})} = \frac{g(\underline{\theta},\ell(\underline{r}))h(\underline{r})f(\underline{\theta})}{f(\underline{r})}$$

Now, $f(\underline{r}) = \int f(\underline{r}|\underline{\theta})f(\underline{\theta})d\underline{\theta}$

$$= h(\underline{r})\int g(\underline{\theta},\ell(\underline{r}))f(\underline{\theta})d\underline{\theta}.$$

Thus,

$$f(\underline{\theta}|\underline{r}) = \frac{g(\underline{\theta},\ell(\underline{r}))f(\underline{\theta})}{\int g(\underline{\theta}',\ell(\underline{r}))f(\underline{\theta}')d\underline{\theta}'}$$

depends only on $\ell(\underline{r})$. Therefore,

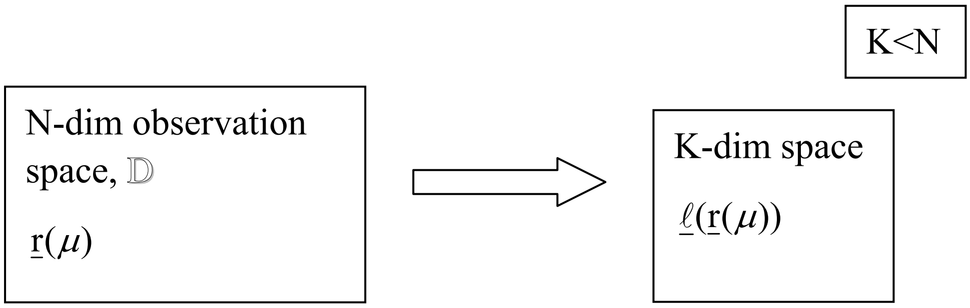
$f(\theta|\underline{r}) = f(\underline{\theta}|\ell(\underline{r}))$ and thus $\ell(\underline{r}(\mu))$ is a sufficient statistic.

"$\Rightarrow$" If $\ell(\underline{r}(\mu))$ is sufficient, then

$$f(\underline{r}|\underline{\theta}) = \frac{f(\underline{\theta}|\underline{r})f(\underline{r})}{f(\underline{\theta})} = \frac{f(\underline{\theta}|\ell(\underline{r}))}{f(\underline{\theta})}f(\underline{r}) = g(\underline{\theta},\ell(\underline{r}))h(\underline{r}). \qquad \text{Q.E.D.}$$

Notes: (1) The above statements can be restated for a set of sufficient statistics, $\underline{\ell}(\underline{\mathrm{r}}(\mu))$, and all properties hold true.

(2) Let $\underline{\ell}(\underline{\mathrm{r}}(\mu))$ be a *K*-dimensional sufficient statistic with $\underline{\mathrm{r}}(\mu)$ being *N*-dimensional and *K*<*N*. Then, $\underline{\ell}(\underline{\mathrm{r}}(\mu))$ transforms the *N*-dimensional observation space into a smaller *K*-dimensional sufficient space, without losing any information on $\underline{\theta}$.



(3) Sufficient statistics are not unique because $a\underline{\ell}(\underline{\mathrm{r}}(\mu))$ is also a sufficient statistic for any *a*.

*Minimum Error Probability (MEP) Criterion

•Define: $\mathrm{P_e} \equiv$ Error Probability

$$\begin{aligned}
&\equiv \Pr\{\text{"H}_1 \text{ is sent and H}_0 \text{ is detected"} \\
&\qquad\qquad \text{or "H}_0 \text{ is sent and H}_1 \text{ is detected"}\} \\
&= \Pr\{\text{"H}_1 \text{ is sent" and "H}_0 \text{ is detected"}\} \\
&\qquad\qquad + \Pr\{\text{ "H}_0 \text{ is sent" and "H}_1 \text{ is detected"}\} \\
&= P(\mathrm{H}_1)\Pr\{\mathrm{H}_0 \text{ is detected}|\mathrm{H}_1 \text{ is sent}\} \\
&\qquad\qquad + P(\mathrm{H}_0)\Pr\{\mathrm{H}_1 \text{ is detected}|\mathrm{H}_0 \text{ is sent}\} \\
&= P(\mathrm{H}_1)\int_{D_0} f(\underline{\mathrm{r}}|\mathrm{H}_1)d\underline{\mathrm{r}} + P(\mathrm{H}_0)\int_{D_1} f(\underline{\mathrm{r}}|\mathrm{H}_0)d\underline{\mathrm{r}}
\end{aligned}$$

•If we set $\mathrm{C}_{00} = \mathrm{C}_{11} = 0$ (No cost for correct decision)

and $\mathrm{C}_{01} = \mathrm{C}_{10} = 1$ (Unit or equal cost for erroneous decision),

then $\overline{\mathrm{C}} = \mathrm{P_e}$.

Thus, MEP Criterion $\equiv$ Bayes Criterion with $\mathrm{C}_{00} = \mathrm{C}_{11} = 0$ and $\mathrm{C}_{01} = \mathrm{C}_{10} = 1$.

From previous discussion on Bayes Criterion, MEP criterion leads us to the test

$$\Lambda(\underline{\mathrm{r}}) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta_{\mathrm{MEP}} \quad \text{with } \eta_{\mathrm{MEP}} \equiv \frac{\mathrm{P(H_0)}}{\mathrm{P(H_1)}}$$

Notes:

1. Bayes rule requires the cost function and *a priori* probability for threshold implementation.
2. MEP rule requires the *a priori* probability in implementation.
3. In digital communications, *a priori* probabilities are usually feasible (Equal *a priori* probabilities are frequently achieved for unconstrained channels.)
4. $$\frac{\Lambda(\underline{r})}{\eta_{MEP}} = \frac{f(\underline{r}|H_1)P(H_1)}{f(\underline{r}|H_0)P(H_0)} = \frac{P(H_1|\underline{r})f(\underline{r})}{P(H_0|\underline{r})f(\underline{r})} = \frac{P(H_1|\underline{r})}{P(H_0|\underline{r})} \overset{H_1}{\underset{H_0}{\gtrless}} 1 \quad \cdots\cdots\cdots \oplus$$

   The MEP test is to find H_i that gives the maximum *a posteriori* probability (MAP). This rule $\oplus$ is called an MAP rule.

*Further Terminology

$$P_F \equiv \text{Probability of false alarm}$$
$$\equiv \Pr\{H_1 \text{ is decided}|H_0 \text{ is sent}\}$$
$$= \int_{D_1} f(\underline{r}|H_0)d\underline{r}$$

$$P_D \equiv \text{Probability of detection}$$
$$\equiv \Pr\{H_1 \text{ is decided}|H_1 \text{ is sent}\} = \int_{D_1} f(\underline{r}|H_1)d\underline{r}$$

$$P_M \equiv \text{Probability of miss detection}$$
$$\equiv \Pr\{H_0 \text{ is decided}|H_1 \text{ is sent}\}$$
$$= \int_{D_0} f(\underline{r}|H_1)d\underline{r}$$
$$= 1 - P_D$$

*Minimax Criterion

The test is to minimize the maximum possible average cost, in the absence of a priori probabilities.

Now, recall that

$$\overline{C} = P(H_0)C_{10} + P(H_1)C_{11} + P(H_1)(C_{01}\text{-}C_{11})P_M \text{-} P(H_0)(C_{10}\text{-}C_{00})(1\text{-}P_F)$$
$$= C_{00}(1-P_F) + C_{10}P_F + P(H_1)\{(C_{11}\text{-}C_{00}) + (C_{01}\text{-}C_{11})P_M \text{-}(C_{10}\text{-}C_{00})P_F\}$$

since $P(H_0) = 1\text{-}P(H_1)$. The minimax criterion can be mathematically expressed as

$$\min_{D_0} \max_{P(H_1)} \overline{C} = \min_{D_0} \{C_{00}(1-P_F) + C_{10}P_F + \max_{P(H_1)} P(H_1)\{\underbrace{(C_{11}\text{-}C_{00}) + (C_{01}\text{-}C_{11})P_M \text{-}(C_{10}\text{-}C_{00})P_F}_{\equiv\Delta}\}\}$$

If $\Delta > 0$,

$$\max_{P(H_1)} P(H_1)\Delta = \Delta \text{ and}$$

$$\text{Criterion} = \min_{D_0} \{C_{00}(1-P_F) + C_{10}P_F + \Delta\}$$
$$= \min_{D_0} \{(C_{01} - C_{11})P_M\} + C_{11}$$
$$= \min\{ \underbrace{C_{11}}_{\substack{D_0=\phi \\ C_{01}\geq C_{11}}} , \underbrace{C_{01}}_{\substack{D_1=\phi \\ C_{01}<C_{11}}} \}$$

If $\Delta < 0$,

$$\max_{P(H_1)} P(H_1)\Delta = 0 \text{ and}$$

$$\text{Criterion} = C_{00} + \min_{D_0} P_F(C_{10}\text{-}C_{00})$$
$$= \min\{ \underbrace{C_{00}}_{\substack{D_1=\phi \\ C_{10}\geq C_{00}}} , \underbrace{C_{10}}_{\substack{D_0=\phi \\ C_{10}<C_{00}}} \}$$

Both cases yield “no detection.”

Thus, the optimization can be obtained when

$\Delta = 0$, i.e.,

$$(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_F = 0$$

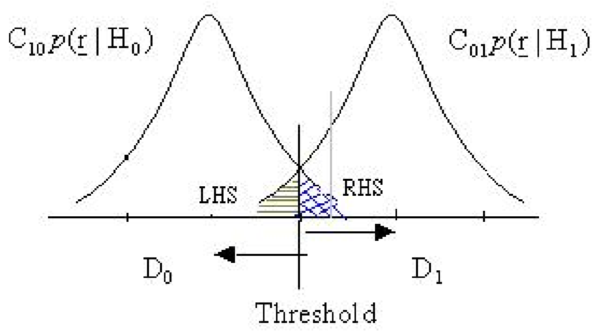
which is called the minimax equation.

• Special Case $(C_{00} = C_{11} = 0)$

The minimax equation becomes

$$C_{01}P_M = C_{10}P_F$$

$$\Rightarrow \underbrace{\int_{D_0} C_{01} f(\underline{r} \mid H_1) d\underline{r}}_{LHS} = \underbrace{\int_{D_1} C_{10} f(\underline{r} \mid H_0) d\underline{r}}_{RHS} \quad \cdots\cdots \otimes$$



If threshold is increased, LHS goes up and RHS goes down.

If threshold is decreased, LHS goes down and RHS goes up.

There is a threshold satisfying $\otimes$

Note: Minimax test requires the cost function in implementation

* Neyman-Pearson (NP) Criterion

The criterian is to maximize P_D (or minimize P_M) under the constraint that $P_F \le \alpha$, where $\alpha < 1$ is a design value.

This is a constrained optimization problem, which can be solved by using Lagrange multipliers. That is, the criterian is to do the following:

$$\min_{D_0} P_M + \lambda[P_F - \alpha],\ \lambda > 0$$

where λ is the Lagrange multiplier.

$$\Rightarrow \min_{D_0} \int_{D_0} f(\underline{r}|H_1)d\underline{r} + \lambda[\int_{D_1} f(\underline{r}|H_0)d\underline{r} - \alpha]$$

$$\Rightarrow \min_{D_0} \int_{D_0} [f(\underline{r}|H_1) - \lambda f(\underline{r}|H_0)]d\underline{r} + \lambda(1-\alpha)$$

For any $\lambda > 0$, choosing D_0 such that

$f(\underline{r}\,|\,H_1) - \lambda f(\underline{r}\,|\,H_0) < 0$

will achieve the optimization. Thus, the NP test has the form of LRT

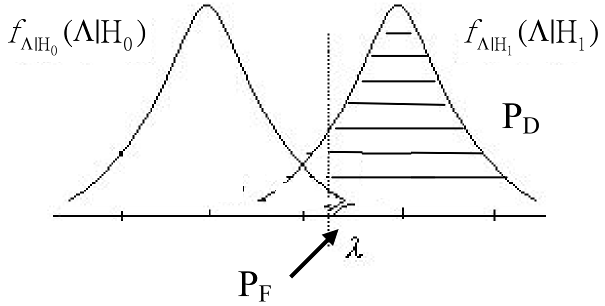
$$\Lambda(\underline{r}) = \frac{f(\underline{r}\,|\,H_1)}{f(\underline{r}\,|\,H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

To satisfy the constraint, λ has to be chosen in a way that $P_F \le \alpha$, i.e.

$$P_F = \int_{\lambda}^{\infty} f(\Lambda\,|\,H_0)d\Lambda \le \alpha$$

Notes:

1. In most of cases, $f(\Lambda|H_i)$ is continuous. Thus, the above equality should be used since



2. NP test does not need cost function nor a priori probabilities in implementation.
3. If $\lambda < 0$, then the criterion is

$$\min_{D_0} P_M + \lambda[\alpha - P_F], \;\; \lambda < 0$$

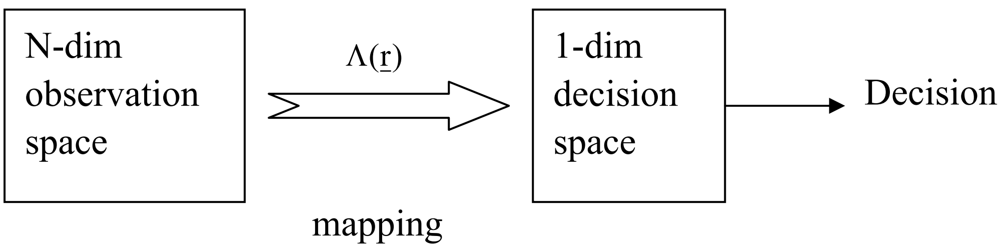
4. NP test is frequently used in radar and sonar detections.

*Observation

1. The tests based on Bayes, MEP, and NP criteria consist of LRT, namely

$$\Lambda(\underline{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad \cdots\cdots \oplus$$

2. The LRT implies the decision process



That is, the decision space is 1-dim, regardless of the dimensionality of the observation space.

3. Instead of the observed statistic $\underline{r}(\mu)$, a statistic $\Lambda(\underline{r}(\mu))$ is sufficient for the LRT purpose.

Now, define the transformation

$$(\mathrm{r}_1,\mathrm{r}_2,\cdots,\mathrm{r}_N) \rightarrow (\ell,\underbrace{\mathrm{y}_1,\mathrm{y}_2,\cdots,\mathrm{y}_{N\text{-}1}}_{\underline{\mathrm{y}}})$$

where $\ell(\mu)$ is a sufficient statistic for the test $\oplus$.

Then, we can express the $\Lambda(\underline{\mathrm{r}})$ as

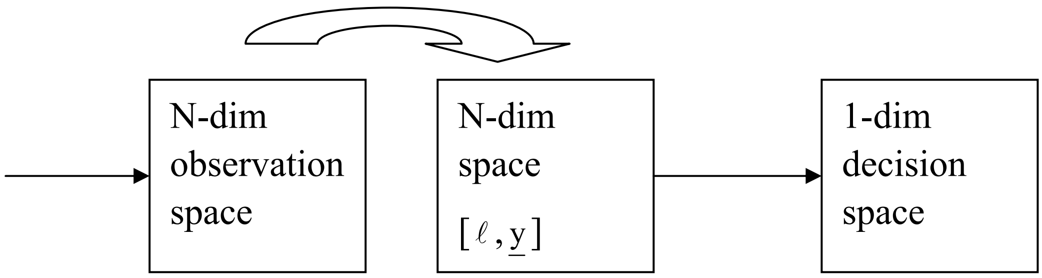
$$\Lambda(\underline{\mathrm{r}}) = \Lambda(\ell,\underline{\mathrm{y}}) = \frac{f(\ell,\underline{\mathrm{y}}|\mathrm{H}_1)}{f(\ell,\underline{\mathrm{y}}|\mathrm{H}_0)}$$

$$\Rightarrow \Lambda(\ell,\underline{\mathrm{y}}) = \frac{f(\ell|\mathrm{H}_1)f(\underline{\mathrm{y}}|\ell,\mathrm{H}_1)}{f(\ell|\mathrm{H}_0)f(\underline{\mathrm{y}}|\ell,\mathrm{H}_0)}$$

Since $\ell(\mu)$ is a sufficient statistic for the test $\oplus$

$$\Lambda(\underline{\mathrm{r}}) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta \text{ is equivalent to } \Lambda(\ell) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta \text{ with } \Lambda(\ell) = \frac{f(\ell|\mathrm{H}_1)}{f(\ell|\mathrm{H}_0)}$$

This implies that $f(\underline{\mathrm{y}}|\ell,\mathrm{H}_1) = f(\underline{\mathrm{y}}|\ell,\mathrm{H}_0)$. That is, in the new coordinate system $(\ell,\mathrm{y}_1,\mathrm{y}_2,\cdots,\mathrm{y}_{N\text{-}1})$, the first coordinate (the one denoted by ℓ) contains all the information necessary for the decision problem.



Example: (Ex* with N = 2)

H_0: $r_1(\mu) = n_1(\mu)$

$r_2(\mu) = n_2(\mu)$ with $n_1(\mu)$ and $n_2(\mu)$ i.i.d. and $\sim G(0,\sigma^2)$

H_1: $r_1(\mu) = m + n_1(\mu)$

$r_2(\mu) = m + n_2(\mu)$

From Ex*,

$$\Lambda(\underline{r}) = \prod_{i=1}^{2} e^{\frac{-1}{2\sigma^2}(m^2 - 2r_i m)}$$

$$\underbrace{\Rightarrow}^{\text{log LRT}} \quad \ell(\underline{r}) = (\sum_{i=1}^{2} r_i) \cdot \text{constant}$$

Now, define $(r_1, r_2) \rightarrow (\ell, y)$ with

$$\ell = (r_1 + r_2)/\sqrt{2}$$

$$y = (r_1 - r_2)/\sqrt{2}$$

$\Rightarrow f(y|H_0) \sim G(0,\sigma^2)$

$f(y|H_1) \sim G(0,\sigma^2)$

$\Rightarrow f(y|H_0) = f(y|H_1)$

Also, $\ell(\mu)$ and $y(\mu)$ are independent because $\ell(\mu)$ and $y(\mu)$ are uncorrelated and jointly Gaussian. Thus, $f(y|\ell, H_i) = f(y|H_i)$. This gives $f(y|\ell, H_1) = f(y|\ell, H_0)$.