Session 8

8.1

Detection Theory

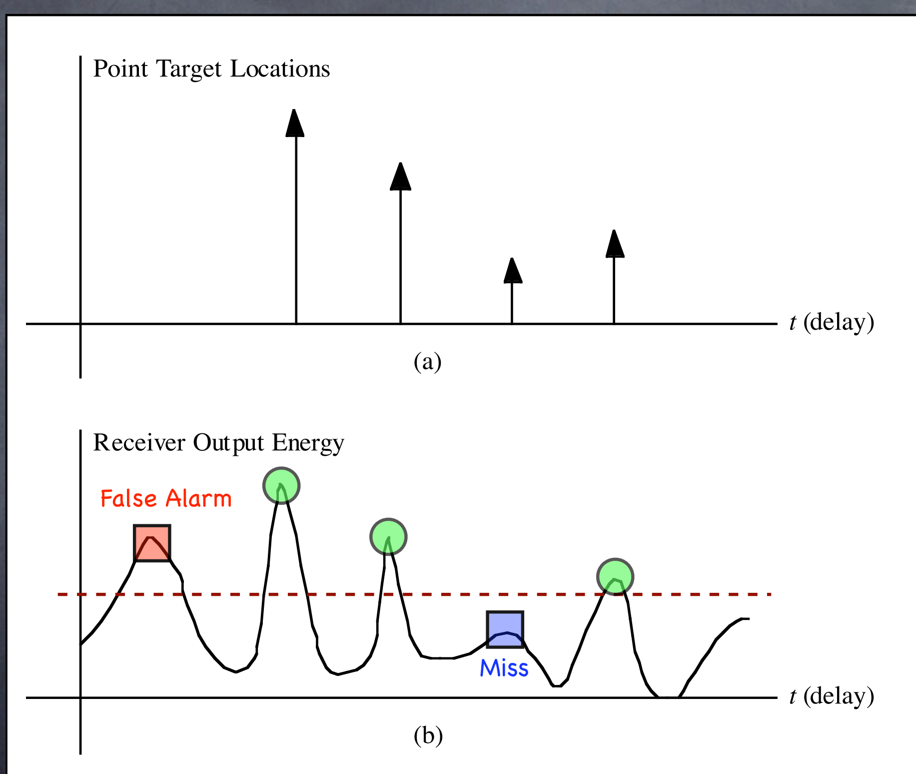
- The most basic task of a RADAR system is to detect whether or not an object (target) is present.
- We may also wish to detect if some "feature" is present in the scattered signal:
  - Spectral Component
  - signal family characteristic
  - depolarized wave component –

8.2

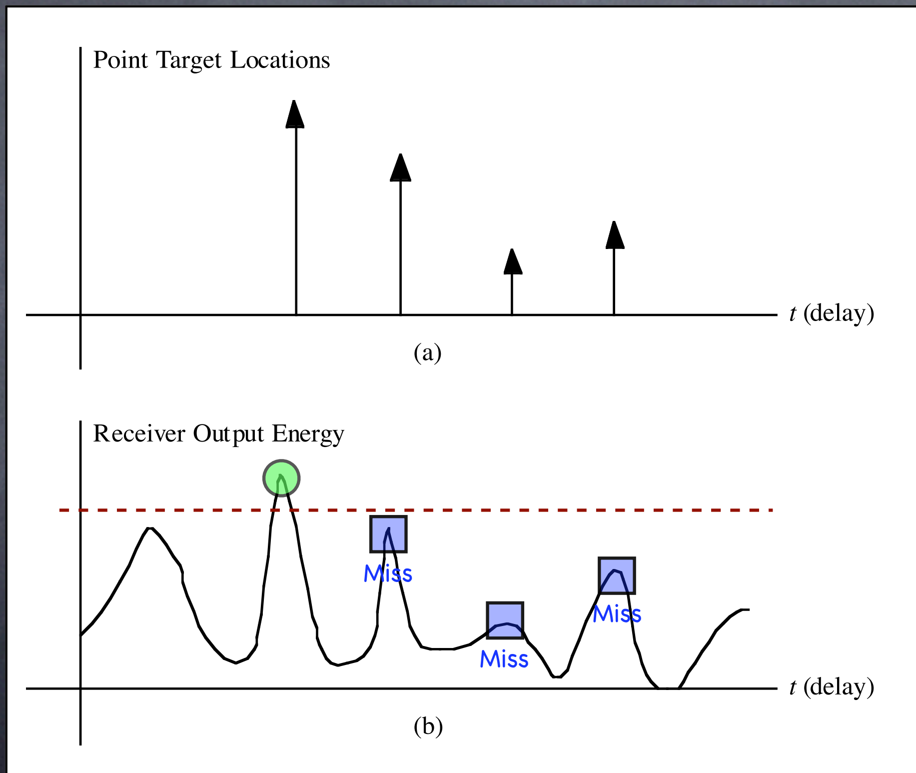
- This almost always involves making a decision in the presence of noise.
- Noise can generally—at best—be described statistically (probabilistically.)
- Statistical Decision Theory/Hypothesis Testing is the tool of choice for these problems.
- Engineers usually call this area of statistics Detection Theory.

8.3

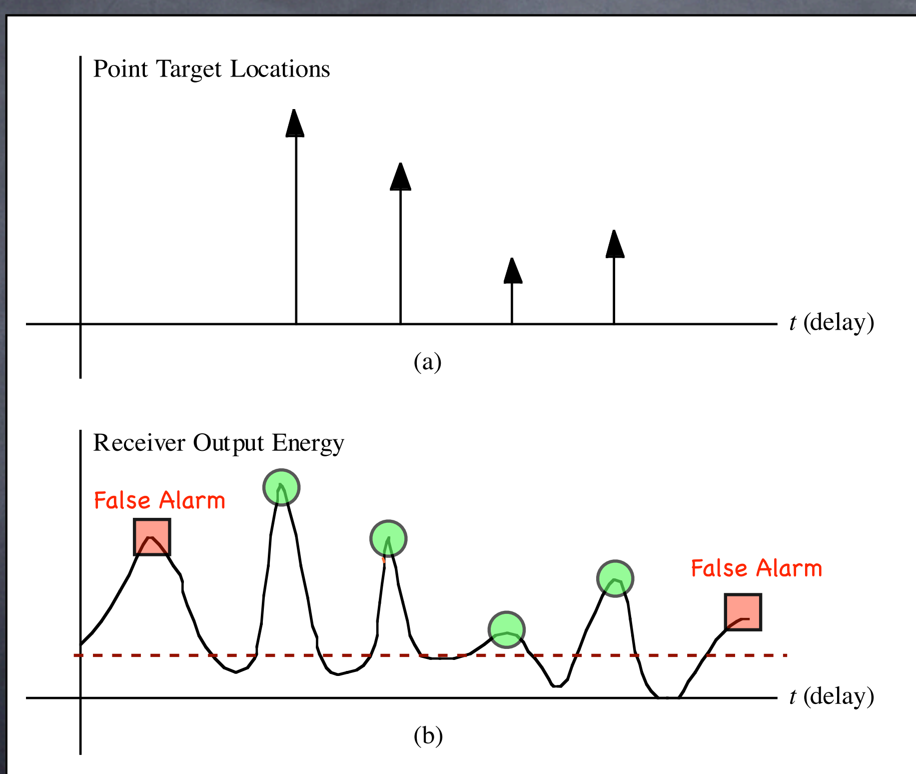
Example: Detection Based on Energy (A-Scope)



Don't Like False Alarms: Raise the Threshold



Don't Like Misses: Lower the Threshold



8.6

An Optimal Test?

- How do we design an optimal test in this case?
- We want as small a number of "false alarms" as possible.
- We want as small a number of "misses" as possible.
- Clearly theses two goals contradict each other.
- An "optimal" test must trade these two goals off against each other. But how?

8.7

In general, we deal with observations from our sensing system of the form

$$y(t) = y[s(t), n(t), \underline{\alpha}(t)]$$

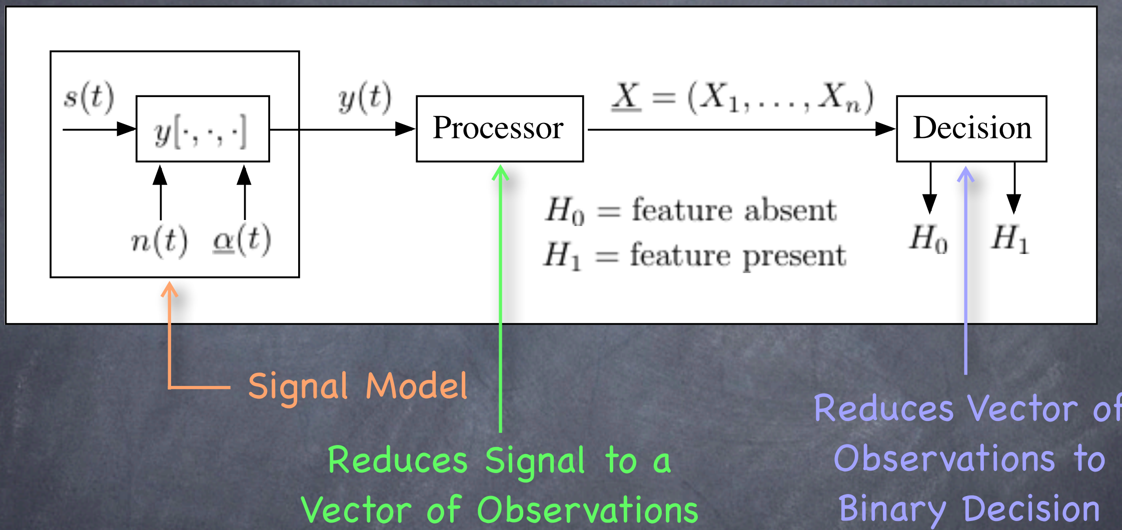
where

$s(t)$ = transmitted signal

$n(t)$ = noise in received signal

$\underline{\alpha}(t)$ = "state" of the *system* (radar, target, medium)

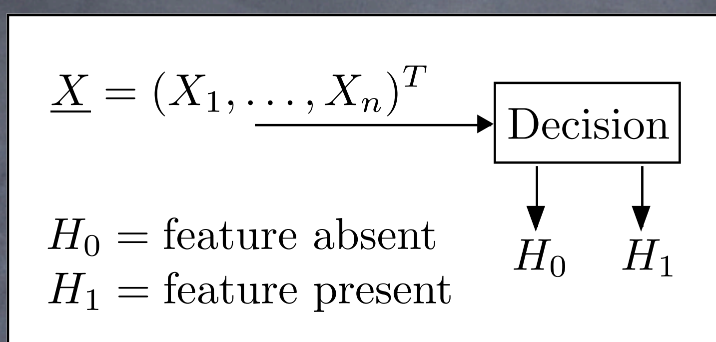
General Framework for Radar Detection



Input to processor is a random process $y(t)$.

Processor output $\underline{X} = (X_1, \ldots, X_N)^T$ is a vector of jointly distributed random variables—a *random vector*.

For now, we will focus on the decision block



We will assume that after processing by the processor, the the random vector $\underline{X}$ is described by a cdf parameterized a parameter vector $\underline{\theta}$:

$$F_{\underline{\theta}}(\underline{x}) = \mathrm{P}_{\underline{\theta}}(\{X_1 \le x_1\} \cap \cdots \cap \{X_N \le x_N\})$$

How do we design optimal decision blocks?

8.10

Binary Hypothesis Testing

Let

$$\underline{X} = (X_1, \ldots, X_N)^T = \text{vector of observations}$$

governed by

$$F_{\underline{\theta}}(\underline{x}) = \mathrm{P}_{\underline{\theta}}(\{X_1 \le x_1\} \cap \cdots \cap \{X_N \le x_N\}),$$

where

$$\mathrm{P}_{\underline{\theta}} = \text{probability measure parameterized by } \underline{\theta}$$

8.11

For example:

Suppose $X_1, \ldots, X_N$ are i.i.d Gaussians with mean μ and variance σ^2.

$$\begin{aligned} F_{\underline{\theta}}(\underline{x}) &= F_{\underline{\theta}}(x_1, \ldots, x_N) \\ &= F_{\underline{\theta}}(x_1) F_{\underline{\theta}}(x_2) \cdots F_{\underline{\theta}}(x_N) \\ &= \Phi\left(\frac{x_1 - \mu}{\sigma}\right) \Phi\left(\frac{x_2 - \mu}{\sigma}\right) \cdots \Phi\left(\frac{x_n - \mu}{\sigma}\right) \end{aligned}$$

where

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-v^2/2}\, dv,$$

and

$$\underline{\theta} = (\mu, \sigma^2)^T \in \Theta = \mathbf{R} \times [0, \infty)$$

Let Θ_0 and Θ_1 be a partition of Θ (the parameter space) (i.e., $\Theta_0 \cap \Theta_1 = \emptyset$ and $\Theta_0 \cup \Theta_1 = \Theta$)

Either

$$\underline{\theta} \in \Theta_0 \qquad \Rightarrow \qquad H_0 \text{ is true}$$

or

$$\underline{\theta} \in \Theta_1 \qquad \Rightarrow \qquad H_1 \text{ is true}$$

We will assume both H_0 and H_1 are *simple hypotheses*.

Defn. A hypothesis H_i is said to be a *simple hypothesis* if its corresponding parameter set Θ_i contains only a single element $\underline{\theta}_i$. If Θ_i contains more than one element, then H_i is a *composite hypothesis*.

If H_0 and H_1 are both simple, then

$$\Theta_0 = \{\underline{\theta}_0\} \text{ and } \Theta_1 = \{\underline{\theta}_1\}$$

and

$$\Theta = \Theta_0 \cup \Theta_1 = \{\underline{\theta}_0, \underline{\theta}_1\}$$

Then if hypothesis H_i is true, $\underline{X}$ has a single cdf $F_{\underline{\theta}_i}(\underline{x})$, for $i = 0, 1$.

For a composite H_i, there is a corresponding set of cdfs:

$$\{F_{\underline{\theta}}(\underline{x}); \underline{\theta} \in \Theta_i\}$$

Hypothesis H_i is true if any one of these cdfs describes $\underline{X}$.

Given a random vector $\underline{X}$ having cdf $F_{\underline{\theta}}(\underline{x})$, $\underline{\theta} \in \Theta$, we can define a binary test of the form

$$\phi(\underline{x}) = \begin{cases} 1, & \text{for } \underline{x} \in \mathcal{R}, \\ 0, & \text{for } \underline{x} \in \mathcal{A}, \end{cases}$$

where

$$\mathcal{R} = \textit{rejection region} \text{ of } H_0$$

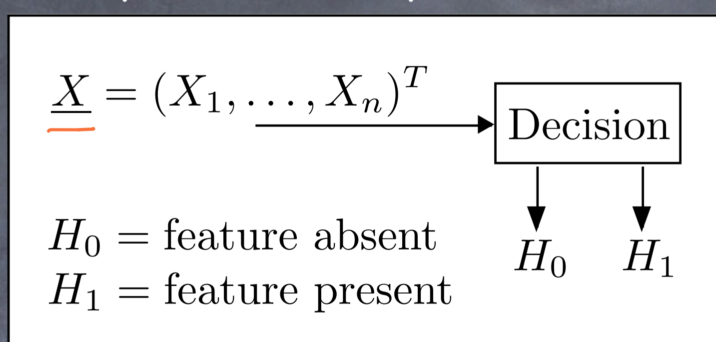
$$\mathcal{A} = \textit{acceptance region} \text{ of } H_0$$

Assuming real RVs $X_1, \ldots, X_N$, $\underline{x} \in \mathbf{R}^N$, and

$$\mathcal{R} \subset \mathbf{R}^N \qquad \mathcal{R} \cap \mathcal{A} = \emptyset$$

$$\mathcal{A} \subset \mathbf{R}^N \qquad \mathcal{R} \cup \mathcal{A} = \mathbf{R}^N$$

Recap: Binary Hypothesis tests



$$\underline{X} = (X_1, \ldots, X_N)^T = \text{vector of observations}$$

governed by

$$F_{\underline{\theta}}(\underline{x}) = \mathrm{P}_{\underline{\theta}}(\{X_1 \le x_1\} \cap \cdots \cap \{X_N \le x_N\}),$$

We will assume both H_0 and H_1 are *simple hypotheses*.

$$\text{Under } H_0,\ \Theta_0 = \{\underline{\theta}_0\} \Rightarrow \underline{X} \sim F_{\underline{\theta}_0}(\underline{x})$$

$$\text{Under } H_1,\ \Theta_1 = \{\underline{\theta}_1\} \Rightarrow \underline{X} \sim F_{\underline{\theta}_1}(\underline{x})$$

8.16

Given a random vector $\underline{X}$ having cdf $F_{\underline{\theta}}(\underline{x})$, $\underline{\theta} \in \Theta$, we can define a binary test of the form

$$\phi(\underline{x}) = \begin{cases} 1, & \text{for } \underline{x} \in \mathcal{R}, \\ 0, & \text{for } \underline{x} \in \mathcal{A}, \end{cases}$$

where

$$\mathcal{R} = \textit{rejection region} \text{ of } H_0$$

$$\mathcal{A} = \textit{acceptance region} \text{ of } H_0$$

Assuming real RVs $X_1, \ldots, X_N$, $\underline{x} \in \mathbf{R}^N$, and

$$\mathcal{R} \subset \mathbf{R}^N \qquad \mathcal{R} \cap \mathcal{A} = \emptyset$$

$$\mathcal{A} \subset \mathbf{R}^N \qquad \mathcal{R} \cup \mathcal{A} = \mathbf{R}^N$$

8.17

$$\phi(\underline{x}) = \begin{cases} 1, & \text{for } \underline{x} \in \mathcal{R}, \\ 0, & \text{for } \underline{x} \in \mathcal{A}, \end{cases}$$

$$\mathcal{R} = \textit{rejection region} \text{ of } H_0$$

$$\mathcal{A} = \textit{acceptance region} \text{ of } H_0$$

Note the language:

If $\phi(\underline{X}) = 0$, we accept hypothesis H_0.

If $\phi(\underline{X}) = 1$, we reject hypothesis H_0.

~~(and effectively accept H_1)~~

In general, two types of errors can occur:

H_0 is true, but $\phi(\underline{X}) = 1$. *Type I Error* or *False Alarm*

H_1 is true, but $\phi(\underline{X}) = 0$. *Type II Error* or *Miss*

The probability of Type I error is called the *size* or *probability of false alarm* of the test.

It is given by

$$\alpha = P_{\underline{\theta}_0}(\{\phi(\underline{X}) = 1\})$$

when H_0 is simple.

Note that the size can also be written as

$$\alpha = P_{\underline{\theta}_0}(\{\phi(\underline{X}) = 1\})$$

$$= \boxed{E_{\underline{\theta}_0}[\phi(\underline{X})]}$$

n.b. the subscripts $\underline{\theta}_0$ mean these quantities are computed using cdf $F_{\underline{\theta}_0}(\underline{x})$ or corresponding pdf $f_{\underline{\theta}_0}(\underline{x})$

If H_0 is composite, the size is defined as

$$\alpha = \sup_{\underline{\theta} \in \Theta_0} \left\{ E_{\underline{\theta}}[\phi(\underline{X})] \right\}$$

Here "sup" means *supremum* (i.e., the least upper bound)

©2004 by Mark R. Bell, mrb@ecn.purdue.edu

8.20

Note that

$$\alpha = \sup_{\underline{\theta} \in \Theta_0} \left\{ \mathrm{E}_{\underline{\theta}}[\phi(\underline{X})] \right\}$$

represents the worst-case false alarm for any possible $\underline{\theta}_0 \in \Theta_0$.

If H_1 is in effect and $\phi(\underline{X}) = 1 \;\Rightarrow\;$ Correct Decision

If H_1 is simple, the probability of correctly deciding H_1 is called the *power* or *probability of detection* of $\phi(\cdot)$:

$$\beta = P_{\underline{\theta}_1}(\{\phi(\underline{X}) = 1\}) = \mathrm{E}_{\underline{\theta}_1}[\phi(\underline{X})]$$

8.21

If H_1 is composite, we compute β for each possible $\underline{\theta}_1 \in \Theta_1$:

$$\beta(\underline{\theta}) = P_{\underline{\theta}}(\{\phi(\underline{X}) = 1\}) = \mathrm{E}_{\underline{\theta}}[\phi(\underline{X})], \quad \forall \underline{\theta} \in \Theta_1$$

Note the asymmetry between H_0 and H_1:

$$\alpha = \sup_{\underline{\theta} \in \Theta_0} \left\{ \mathrm{E}_{\underline{\theta}}[\phi(\underline{X})] \right\} \qquad \text{(worst case)}$$

$$\beta(\underline{\theta}) = \mathrm{E}_{\underline{\theta}}[\phi(\underline{X})], \quad \forall \underline{\theta} \in \Theta_1 \qquad \text{(each case)}$$

Why? Fundamental nature of frequentist hypothesis test.

8.22

A Radar Interpretation

$$\alpha = \sup_{\underline{\theta} \in \Theta_0} \left\{ \mathrm{E}_{\underline{\theta}}[\phi(\underline{X})] \right\} \qquad \text{(worst case)}$$

$\Rightarrow \alpha$ is maximum possible probability of false alarm.

$$\beta(\underline{\theta}) = \mathrm{E}_{\underline{\theta}}\left[\phi(\underline{X})\right], \quad \forall \underline{\theta} \in \Theta_1 \qquad \text{(each case)}$$

$\Rightarrow$ Evaluate the probability of detection for each possible target.

8.23

Because β is the "probability of detection," we sometimes write it as

$$P_D = \beta = \text{Probability of Detection}$$

The probability of deciding that H_0 is in effect when in fact H_1 is —a type II error—is called the *miss probability*

$$P_M = 1 - \beta = 1 - P_D$$

Because α is the probability of false alarm," we sometimes write it as

$$P_{FA} = \alpha = \text{Probability of False Alarm}$$

8.24

We typically characterize the performance of a test ϕ using the pair (α, β), or equivalently (P_{FA}, P_D)

We want to design tests with small α and large β.

Definition: If H_0 and H_1 are both simple, then the test ϕ is *the best test of size α* if it has the greatest β among all tests of size α.

($n.b.$ This implies fixing α and making β as large as possible.)

Constructing a test ϕ amounts to selecting the acceptance region $\mathcal{A}$ and the rejection region $\mathcal{R}$.

If both H_0 and H_1 are simple, we can find the most powerful test of size α using the Neyman-Pearson Lemma.

8.25

The Neyman-Pearson Lemma

Neyman-Pearson Lemma: Let $\Theta = \{\underline{\theta}_0, \underline{\theta}_1\}$, and let $F_{\underline{\theta}_0}(\underline{x})$ be the cdf of the random vector $\underline{X}$ under hypothesis H_0 and $F_{\underline{\theta}_1}(\underline{x})$ be its cdf under hypothesis H_1. Assume that the cdfs $F_{\underline{\theta}_i}(\underline{x})$ have corresponding pdfs or pmfs $f_{\underline{\theta}_i}(\underline{x})$, $i = 0, 1$. Then a test of the form

$$\phi(\underline{x}) = \begin{cases} 1, & \text{for } f_{\underline{\theta}_1}(\underline{x}) > k f_{\underline{\theta}_0}(\underline{x}), \\ \gamma, & \text{for } f_{\underline{\theta}_1}(\underline{x}) = k f_{\underline{\theta}_0}(\underline{x}), \\ 0, & \text{for } f_{\underline{\theta}_1}(\underline{x}) < k f_{\underline{\theta}_0}(\underline{x}), \end{cases}$$

for some $k \geq 0$ and some $0 \leq \gamma \leq 1$ is the most powerful test of size α for testing hypothesis H_0: $\underline{\theta} = \underline{\theta}_0$ versus H_1: $\underline{\theta} = \underline{\theta}_1$.

Interpretation of γ

$$\phi(\underline{x}) = \begin{cases} 1, & \text{for } f_{\underline{\theta}_1}(\underline{x}) > k f_{\underline{\theta}_0}(\underline{x}), \\ \gamma, & \text{for } f_{\underline{\theta}_1}(\underline{x}) = k f_{\underline{\theta}_0}(\underline{x}), \\ 0, & \text{for } f_{\underline{\theta}_1}(\underline{x}) < k f_{\underline{\theta}_0}(\underline{x}), \end{cases}$$

When $\phi(\underline{X}) = \gamma$ we randomly select between H_1 and H_0 such that

$$P(\{\text{Decide } H_1\}|\{\phi(\underline{X}) = \gamma\}) = \gamma$$
$$P(\{\text{Decide } H_0\}|\{\phi(\underline{X}) = \gamma\}) = 1 - \gamma$$

($i.e.$, flip a biased coin having probability γ of comming up "Heads". If "Heads" occurs, decide H_1. If "Tails" occurs, decide H_0.)