Session 20

20.1

Sequential Detection

So far, we have deigned detectors based on a fixed number of observations N:

$$\underline{X} = (X_1, \ldots, X_N).$$

Sometimes, a fixed length $N = N_0$ may be inadequate to achieve the desired detection performance.

At other times, a fixed length $N = N_0$ may be far larger than necessary to achieve the desired detector performance.

In radar problems there is generally a cost incurred for increasing N:

- Time-on-target
- Energy Expended
- Computational load

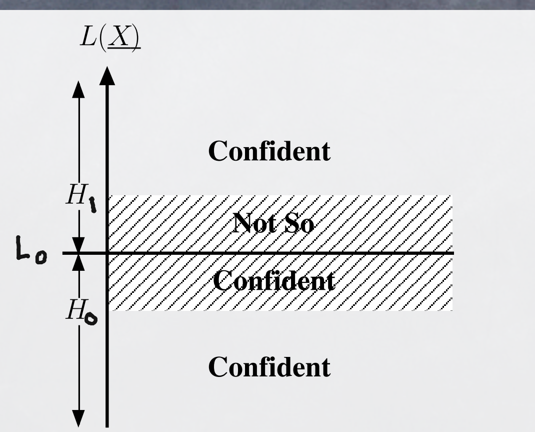
For this reason, we may want to adaptively select the observation length N.

This is the motivation for the idea of *sequential detection*.

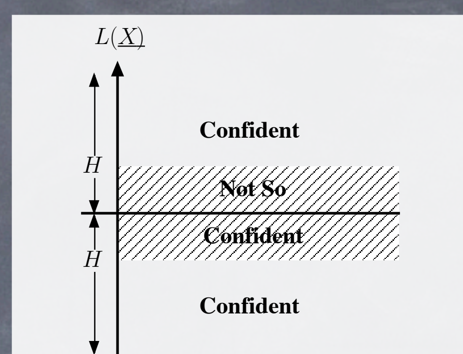
Suppose we make an observation $\underline{X} = (X_1, \ldots, X_N)$ and we use an LRT to decide between two hypotheses H_0 and H_1. Assume our test is of the form

$$L(\underline{X}) = \frac{f_{\underline{\theta}_1}(\underline{X})}{f_{\underline{\theta}_0}(\underline{X})} \underset{H_0}{\overset{H_1}{\gtrless}} L_0.$$

While we would choose H_1 to be true anytime that $L(\underline{X}) > L_0$ and H_0 to be true anytime that $L(\underline{X}) < L_0$, in either case, if $L(\underline{X})$ was very close to L_0, we might not be very confident in our decision.



How can we modify our strategy to get an adaptive length test?



Take a single measurement, look at the resulting likelihood ratio, and make a decision if we can do so with confidence; ...

otherwise, take another measurement ...

20.4

The Sequential Detection Algorithm

1. Let $\underline{X}_1 = (X_1)^T$; set $N = 1$.
2. Calculate $L(\underline{X}_N)$.
3. If $L(\underline{X}_N) > A$, then declare H_1 and stop; else
   If $L(\underline{X}_N) < B$, then declare H_0 and stop; else
   Let $\underline{X}_{N+1} = (\underline{X}_N^T \vdots X_{N+1})^T$, let $N := N + 1$, and go to step 2.

You can make A and B a function of N (i.e., A_N and B_N) but there is usually no reason to do so.

20.5

After N measurements, the likelihood ratio test appears as follows:

$$\phi(\underline{X}_N) = \begin{cases} 1, & \text{for } L(\underline{X}_N) > A, \\ 0, & \text{for } L(\underline{X}_N) < B, \\ \text{take another measurement}, & \text{for } B \leq L(\underline{X}_N) \leq A. \end{cases}$$

In the case where we take another measurement, no decision is made.

If the measurements $X_1, \ldots, X_N, \ldots$ are conditionally independent on the hypotheses H_0 and H_1, then

$$\begin{aligned} L(\underline{X}_N) &= \frac{f_{\underline{\theta}_1}(X_1, \ldots, X_N)}{f_{\underline{\theta}_0}(X_1, \ldots, X_N)} \\ &= \frac{f_{\underline{\theta}_1}(X_N) \cdot f_{\underline{\theta}_1}(X_1, \ldots, X_{N-1})}{f_{\underline{\theta}_0}(X_N) \cdot f_{\underline{\theta}_0}(X_1, \ldots, X_{N-1})} \\ &= L(X_N) \cdot L(\underline{X}_{N-1}) \end{aligned}$$

where $L(X_N)$ is the likelihood ratio based on the single measurement X_N and $L(\underline{X}_{N-1})$ is the likelihood ratio corresponding to the measurement vector $\underline{X}_{N-1}$ made up of the first $N-1$ measurements.

For conditionally independent measurements, the likelihood ratio can be computed recursively!

Sequential Decision Rule

Definition: A *sequential decision rule* $(\underline{\phi}, \underline{\mathcal{S}})$ operates as follows: For an observation sequence $\{X_k; k = 1, 2, \ldots\}$, the rule $(\underline{\phi}, \underline{\mathcal{S}})$ makes a decision $\phi_N(X_1, \ldots, X_N)$, where the random variable N is the *stopping time*, defined by

$$N = \min\{n \in \mathbf{N} : \mathcal{S}_n(X_1, \ldots, X_n) = 1\}.$$

That is, $\mathcal{S}_n(X_1, \ldots, X_n)$ is a *stopping rule* that determines when to stop taking samples. So

$$\mathcal{S}_n(X_1, \ldots, X_n) = \begin{cases} 0, & \text{take another (the } (n+1)\text{-st) sample,} \\ 1, & \text{stop taking samples and make a decision using } \phi_n(\cdot). \end{cases}$$

Sequential Decision Rule (Cont.)

The stopping time N is a random variable, since it depends on the random data sequence. The terminal decision rule $\phi_n(\cdot)$ of $\underline{\phi}$ tells us what decision to make once we stop. An ordinary fixed sample size decision rule using observation vector $\underline{X}_m = (X_1, \ldots, X_m)$ is a special case of a sequential decision rule with

$$\mathcal{S}_n = \begin{cases} 0, & n \neq m, \\ 1, & n = m, \end{cases}$$

and

$$\phi_n(X_1, \ldots, X_n) = \begin{cases} \phi(\underline{X}_n) \text{ the fixed-length test}, & n = m, \\ \text{arbitrary}, & n \neq m. \end{cases}$$

We want to find the optimal sequential Bayes decision rule.

The Optimal Bayes Sequential Test

In order to find the optimal Bayesian sequential test $(\underline{\phi}, \underline{\mathcal{S}})$, we will need to specify both priors and costs.

Assume the priors are $\underline{p} = (p_0, p_1)$ and the costs L_{ij} of deciding hypothesis H_j is in effect when H_i is in fact in effect is

$$L_{ij} = \begin{cases} 1, & \text{when } i \neq j, \\ 0, & \text{when } i = j. \end{cases}$$

We will also assign a cost $C > 0$ to each measurement we make, so that if N is the stop time of our sequential test, the cost of making the measurements is NC.

The assignment of a cost $C > 0$ to each measurement is necessary if we want the test to terminate.

20.10

Bayes Sequential Test (Cont.)

Within this framework, the risks associated with a given sequential test $(\underline{\phi}, \underline{\mathcal{S}})$ are given by

$$R[\theta_0, (\underline{\phi}, \underline{\mathcal{S}})] = \mathrm{E}_{\theta_0}\left[\phi_N(X_1, \ldots, X_N)\right] + \boxed{C \cdot \mathrm{E}_{\theta_0}[N]}$$

Avg. cost of measurements

and

$$R[\theta_1, (\underline{\phi}, \underline{\mathcal{S}})] = 1 - \mathrm{E}_{\theta_1}\left[\phi_N(X_1, \ldots, X_N)\right] + \boxed{C \cdot \mathrm{E}_{\theta_1}[N]}$$

Definition: A *Bayes sequential decision rule* is a sequential decision rule $(\underline{\phi}, \underline{\mathcal{S}})$ that minimizes the Bayes risk

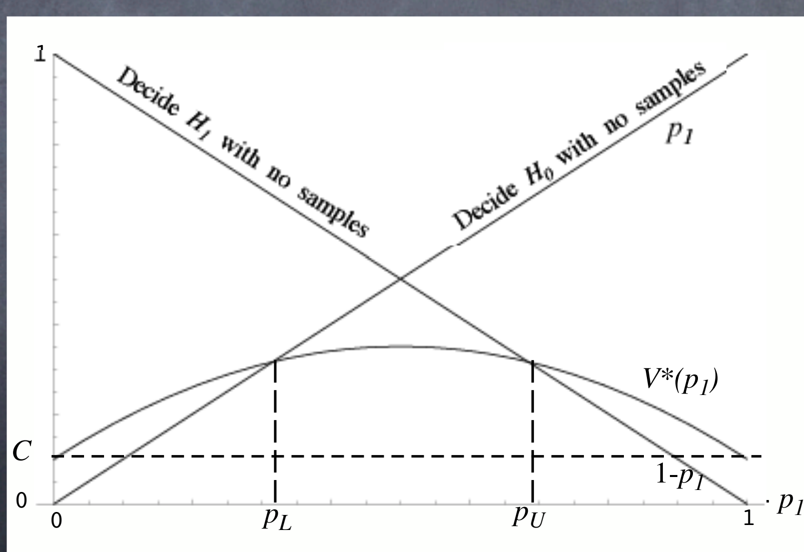
$$R[\mathbf{p}, (\underline{\phi}, \underline{\mathcal{S}})] = p_0 R[\theta_0, (\underline{\phi}, \underline{\mathcal{S}})] + p_1 R[\theta_1, (\underline{\phi}, \underline{\mathcal{S}})].$$

20.11

To see the structure of the optimal sequential Bayes decision rule, consider the function

$$V^*(p_1) = \min_{(\underline{\phi}, \underline{\mathcal{S}})} \left\{ R[(1-p_1, p_1), (\underline{\phi}, \underline{\mathcal{S}})] \right\}$$

where the minimum is taken over all sequential decision rules $(\underline{\phi}, \underline{\mathcal{S}})$ that take at least one sample.



20.12

Let us consider two possible decision rules we can immediately consider:

1. Take no samples and decide H_1. This yields a Bayes risk

$$R[\mathbf{p},,(\underline{\phi},\underline{\mathcal{S}})] = 1 - p_1,$$

where here we have $\phi_0 = \mathcal{S}_0 = 1$.

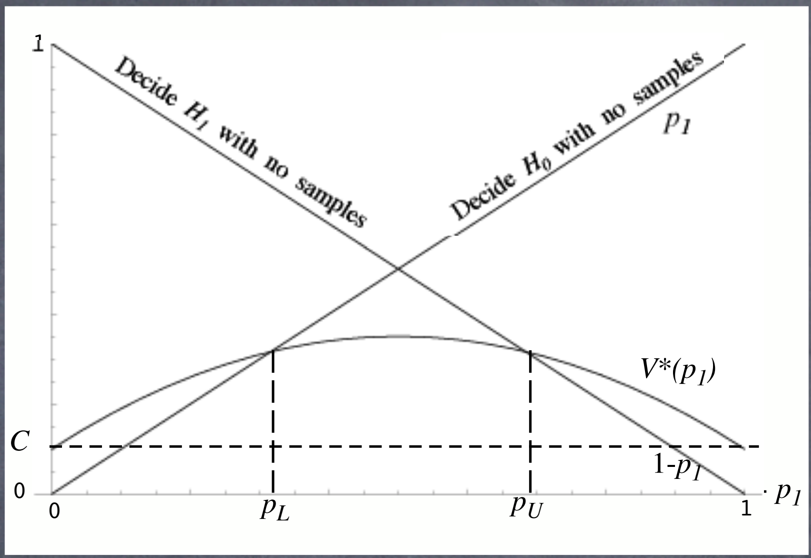
2. Take no samples and decide H_0. This yields a Bayes risk

$$R[\mathbf{p},,(\underline{\phi},\underline{\mathcal{S}})] = p_1,$$

where here we have $\phi_0 = 0$ and $\mathcal{S}_0 = 1$.

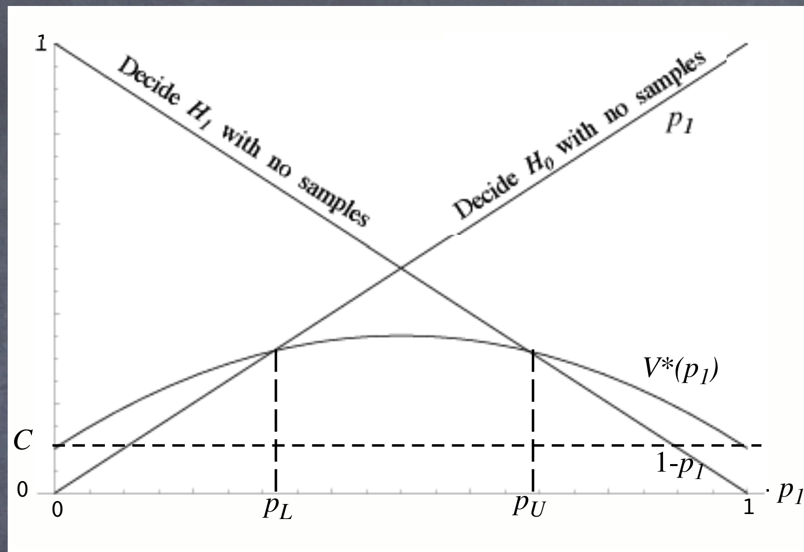
These two rules are not included in the minimization that yielded $V^*(p_1)$, so it may be the case that for some values of p_1 one or the other of these rules may result in a Bayes risk less than $V^*(p_1)$..

20.13



A plot of $V^*(p_1)$ and the Bayes risks $1 - p_1$ and p_1 of the two rules that decide H_1 and H_0, respectively, without taking any samples.

20.14



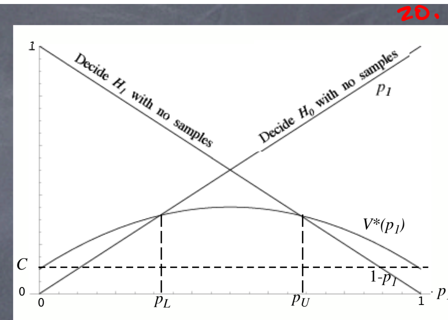
From the figure, we see that:

1. If $p_1 \leq p_L$, then the Bayes sequential test is $\mathcal{S}_0 = 1$ and $\phi_0 = 0$.
2. If $p_1 \geq p_U$, then the Bayes sequential test is $\mathcal{S}_0 = 1$ and $\phi_0 = 1$.
3. If $p_L < p_1 < p_u$, the sequential Bayes test is the sequential decision rule with minimum risk among all $(\underline{\phi}, \underline{\mathcal{S}})$ with $\mathcal{S}_0 = 0$.

20.15

Decision Rule:

1. If $p_1 \leq p_L$, then the Bayes sequential test is $\mathcal{S}_0 = 1$ and $\phi_0 = 0$.
2. If $p_1 \geq p_U$, then the Bayes sequential test is $\mathcal{S}_0 = 1$ and $\phi_0 = 1$.
3. If $p_l < p_1 < p_u$, the sequential Bayes test is the sequential decision rule with minimum risk among all $(\underline{\phi}, \underline{\mathcal{S}})$ with $\mathcal{S}_0 = 0$.



In cases 1 and 2, the test terminates with the corresponding decision because $\mathcal{S}_0 = 1$.

In case 3, the optimal test must take at least one sample, but is otherwise unspecified. Immediately after taking this sample, the probability H_1 is true given the observed sample value is

$$p_1(x_1) = P(\{H_1 \text{ True}\}|\{X_1 = x_1\}).$$

This is our best estimate of the probability that H_1 is true given the information obtained in our first observation.

After taking one sample, the problem of optimizing the test is conditionally the same as before taking any samples in the sense that

1. We still have infinitely many i.i.d. samples available at a cost of C each.
2. All future costs that can be incurred are the same as before we took a sample.

The only difference is that, because we have take one sample, we have more information about which hypothesis is true, and this is reflected in updating our prior p_1 as given by

$$p_1(x_1) = P(\{H_1 \text{ True}\}|\{X_1 = x_1\}).$$

The picture doesn't change—just the prior changes!