Session 21

Recall... 21.1

The Optimal Bayes Sequential Test

In order to find the optimal Bayesian sequential test $(\underline{\phi}, \underline{S})$, 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

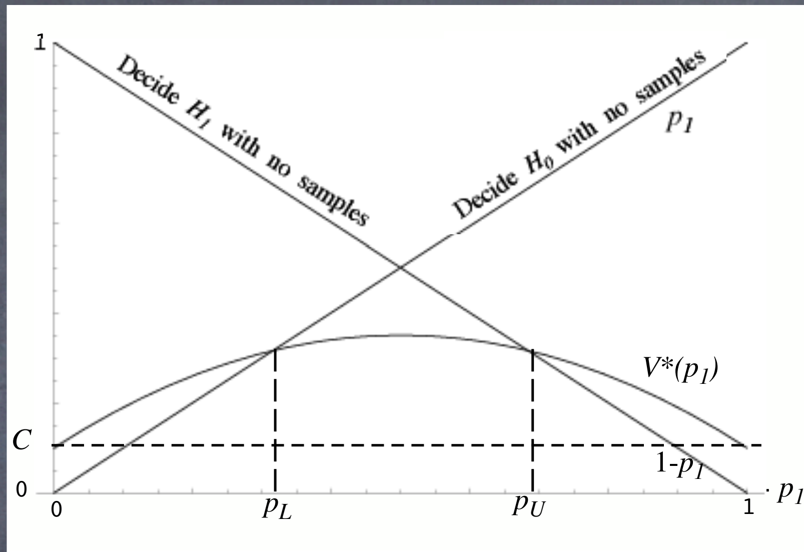
$$L_{ij} = \begin{cases} 1, & \text{when } i \neq j, \\ 0, & \text{when } i = j. \end{cases}$$

Standard Losses

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.

Recall...



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 $\mathcal{S}_0 = 0$.

Recall...

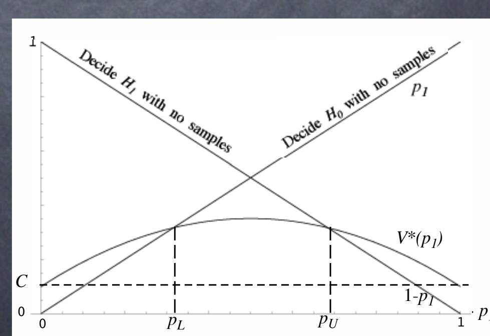
After taking one sample, the problem of optimizing the test is conditionally the same as before taking any samples in the sense that

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!



Thus after taking one sample, the test has the form

1. If $p_1(x_1) \leq p_L$, then the Bayes sequential test is $\mathcal{S}_1 = 1$ and $\phi_1 = 0$.
2. If $p_1(x_1) \geq p_U$, then the Bayes sequential test is $\mathcal{S}_1 = 1$ and $\phi_1 = 1$.
3. If $p_l < p_1(x_1) < p_u$, the sequential Bayes test is the sequential decision rule with minimum risk among all $(\underline{\phi}, \underline{\mathcal{S}})$ with $\mathcal{S}_1 = 0$.

Once again, we either terminate with a decision in case 1 or 2, or we take another sample X_2, update the prior with

$$p_1(x_1, x_2) = p_1(\underline{x}_2) = P(\{H_1 \text{ True}\} | \{X_1 = x_1\} \cap \{X_2 = x_2\}),$$

and repeat the process again.

Continuing with this argument, we have that the Bayes sequential test keeps taking samples until

$$p_1(\underline{x}_n) = P(\{H_1 \text{ True}\} | \{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\}) \notin (p_L, p_U).$$

It then chooses H_0 if $p_1(\underline{x}_N) \leq p_L$ or H_1 if $p_1(\underline{x}_N) \geq p_U$. So the optimal test $(\underline{\phi}, \underline{\mathcal{S}})$ has the stopping rule

$$\mathcal{S}_n(x_1, \ldots, x_n) = \begin{cases} 0, & p_L < p_1(x_1, \ldots, x_n) < p_U, \\ 1, & \text{otherwise,} \end{cases}$$

and the decision rule

$$\phi_n(x_1, \ldots, x_n) = \begin{cases} 1, & p_1(x_1, \ldots, x_n) \geq p_U, \\ 0, & p_1(x_1, \ldots, x_n) \leq p_L. \end{cases}$$

It can be shown that under fairly mild assumptions that with probability one,

$$p_1(x_1, \ldots, x_n) \to 0, \quad \text{as } n \to \infty,$$

under H_0, and

$$p_1(x_1, \ldots, x_n) \to 1, \quad \text{as } n \to \infty,$$

under H_1.

So for any (p_L, p_U) such that $0 < p_L < p_U < 1$, the test eventually converges to a decision with probability one.

All that is needed to specify the optimal Bayes test is p_U and p_L.

Unfortunately, p_U and p_L are not easily calculated except in special cases.

However, the posterior probabilities $p_1(\underline{x}_n)$ are easily calculated.

Assuming the i.i.d. X_k have conditional densities $f_{\underline{\theta}_0}(x)$ and $f_{\underline{\theta}_1}(x)$, we have

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= \frac{p_1 \prod_{k=1}^{n} f_{\underline{\theta}_1}(x_k)}{(1-p_1) \prod_{k=1}^{n} f_{\underline{\theta}_0}(x_k) + p_1 \prod_{k=1}^{n} f_{\underline{\theta}_1}(x_k)} \\ &= \frac{p_1 \lambda_n(x_1, \ldots, x_n)}{p_0 + p_1 \lambda_n(x_1, \ldots, x_n)}, \end{aligned}$$

where

$$\lambda_n(x_1, \ldots, x_n) = \prod_{k=1}^{n} \left(\frac{f_{\underline{\theta}_1}(x_k)}{f_{\underline{\theta}_0(x_k)}} \right) = \lambda_n(x_1, \ldots, x_n) = \left(\frac{f_{\underline{\theta}_1}(x_n)}{f_{\underline{\theta}_0(x_n)}} \right) \lambda_{n-1}(x_1, \ldots, x_{n-1}),$$

where we take $\lambda_0 = 1$ and evaluate recursively with each received sample:

$$\lambda_n(\underline{x}_n) = L(x_n) \lambda_{n-1}(\underline{x}_{n-1}),$$

where

$$\underline{x}_k = (x_1, \ldots, x_k)^T \quad \text{and} \quad L(x_n) = \frac{f_{\underline{\theta}_1}(x_n)}{f_{\underline{\theta}_0}(x_n)}.$$

Such a test is called a *Sequential Probability Ratio Test* (SPRT) for obvious reasons.

21.8

It is in general difficult to analytically determine p_L and p_U for the SPRT, and both *ad hoc* and rigorous techniques. Often *ad hoc* techniques are used to determine values of p_L and p_U that, although they may not be optimal, work well.

We have discussed sequential detection in the Bayesian context, but it can also be used in the classical or frequentist detection framework as well. Here, we will once again find that the SPRT is the optimal test in a certain sense.

For a sequential decision rule $(\underline{\phi}, \underline{\mathcal{S}})$, let $P_{FA}(\underline{\phi}, \underline{\mathcal{S}})$ denote the probability of false alarm (Type I error), and let $P_M(\underline{\phi}, \underline{\mathcal{S}})$ denote the probability of a miss (Type II error), and let $N((\underline{\phi}, \underline{\mathcal{S}}))$ be the stopping time associated with the test. Then if $(\underline{\phi}_0, \underline{\mathcal{S}}_0)$ is the optimal SPRT and $(\underline{\phi}, \underline{\mathcal{S}})$ is any other sequential test (or fixed sample test) and

$$P_{FA}(\underline{\phi}, \underline{\mathcal{S}}) \leq P_{FA}(\underline{\phi}_0, \underline{\mathcal{S}}_0)$$

and

$$P_M(\underline{\phi}, \underline{\mathcal{S}}) \leq P_M(\underline{\phi}_0, \underline{\mathcal{S}}_0),$$

then

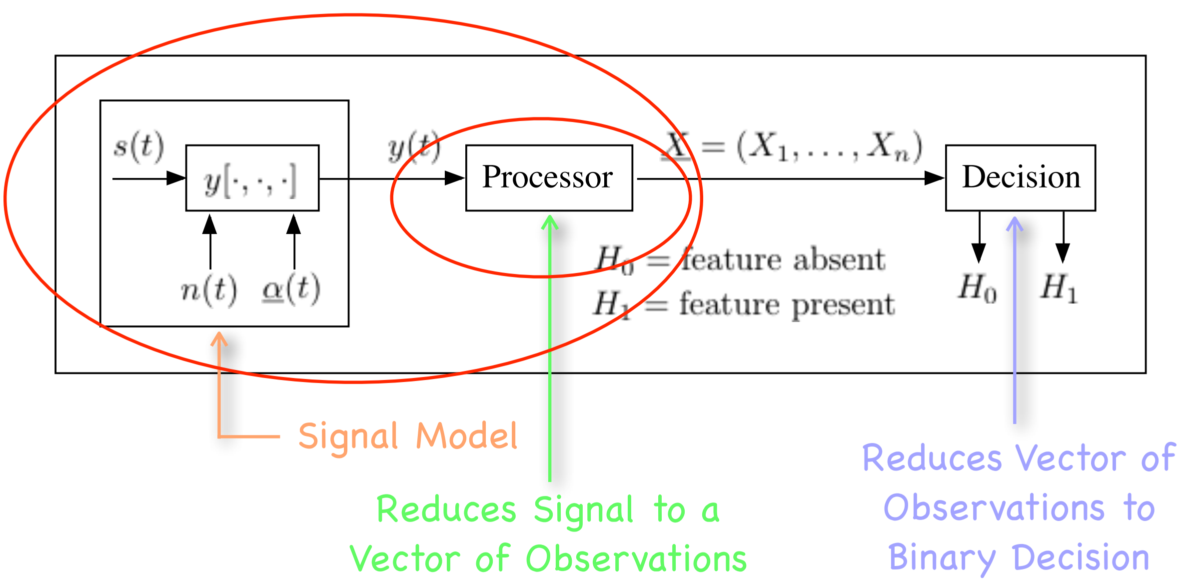
$$\mathrm{E}[N((\underline{\phi}, \underline{\mathcal{S}}))|H_j] \geq \mathrm{E}[N((\underline{\phi}_0, \underline{\mathcal{S}}_0))|H_j], \qquad \text{for } j = 0 \text{ and } 1.$$

Wald-Wolfowitz Theorem

21.9

The Matched Filter

Recall the radar target detection problem:



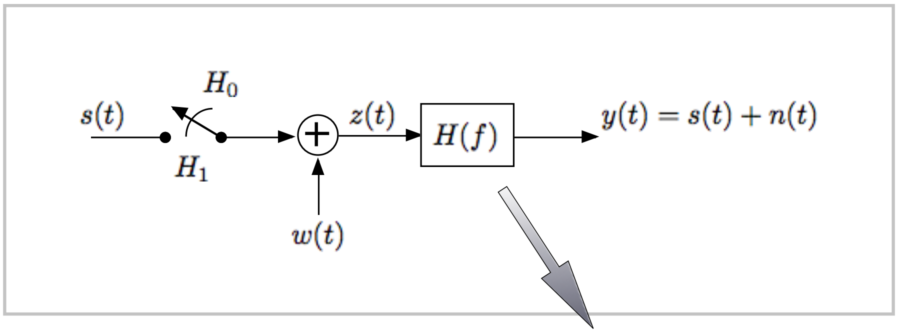
21.10

Detection of a Known Signal in Additive White Gaussian Noise

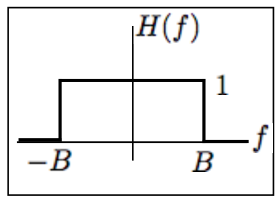
Suppose we have a signal $s(t)$ of known duration T in the interval $[0, T]$ such that

$$s(t) = 0, \quad t \notin [0, T].$$

We wish to determine whether or not this signal is present in the presence of *Additive White Gaussian Noise* (AWGN).



Hypothetical Lowpass Filter:



21.11

Assume the noise $w(t)$ is zero-mean Gaussian white noise having (two-sided) PSD

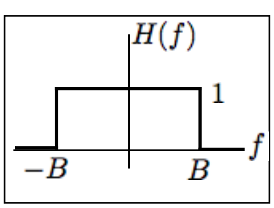
$$S_{ww}(f) = \frac{N_0}{2}. \qquad \left(= \frac{kT_e}{2}\right)$$

We want to determine which of two possible hypotheses are in effect:

$$H_0 : r(t) = w(t) \qquad \text{(target absent)},$$
$$H_1 : r(t) = s(t) + w(t) \qquad \text{(target present)}.$$

Assume we observe the process $z(t)$ through a hypothetical lowpass filter

$$H(f) = 1_{[-B,B]}(f) = \begin{cases} 1, & \text{for } |f| \le B; \\ 0, & \text{for } |f| > B. \end{cases}$$



Assume B sufficiently large such that all but a negligible fraction of the energy in $s(t)$ passes through $H(f)$.

21.12

If we input only the white noise $w(t)$ into filter $H(f)$, the output $n(t)$ becomes bandlimited whote noise $n(t)$:

$$\mathrm{E}[n(t)] = 0$$

$$S_{nn}(f) = \frac{N_0}{2} \cdot 1_{[-B,B]}(f)$$

$$\begin{aligned} R_{nn}(\tau) &= N_0 B \left(\frac{\sin 2\pi B\tau}{2\pi B\tau} \right) \\ &= N_0 B \ \mathrm{sinc}(2B\tau). \end{aligned}$$

It follows that

$$R_{nn}(\tau) = 0, \quad \text{for} \quad \tau = \frac{\pm 1}{2B}, \frac{\pm 2}{2B}, \frac{\pm 3}{2B}, \ldots$$

$$\Longrightarrow \mathrm{E}\left[n\left(t_0 + \frac{k}{2B}\right) \cdot n\left(t_0 + \frac{m}{2B}\right)\right] = N_0 B \delta_{k,m} = \begin{cases} N_0 B, & \text{for } k = m, \\ 0, & \text{for } k \neq m. \end{cases}$$

$$\forall t_0 \in \mathbf{R}$$

21.13

Thus samples of the random process $n(t)$ taken at increments of $\Delta t = 1/(2B)$ form a sequence of uncorrelated Gaussian random variables.

Because this sequence is both Gaussian and uncorrelated, it follows that it is a sequence of independent Gaussian random variables.

Thus $n(t_0 + 1/(2B)), n(t_0 + 2/(2B)), \ldots, n(t_0 + M/(2B))$ are i.i.d. Gaussian random variables with mean zero and variance $\sigma_n^2 = R_{nn}(0) = N_0 B$.

If we take $t_0 = 0$ and sample at time instants $t_m = m/(2B)$, where $m = 1, 2 \ldots 2BT$ over duration T, the p.d.f. under H_0 is

$$\begin{aligned} f_0(y(t_1), \ldots, y(t_{2BT})) &= \prod_{m=1}^{2BT} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left\{ -\frac{y^2(t_m)}{2\sigma_n^2} \right\} \\ &= \frac{1}{(2\pi)^{BT} \sigma_n^{2BT}} \exp\left\{ -\frac{1}{2\sigma_n^2} \sum_{m=1}^{2BT} y^2(t_m) \right\}. \end{aligned}$$

The p.d.f. under H_1 is

$$\begin{aligned} f_1(y(t_1),\ldots,y(t_{2BT})) &= \prod_{m=1}^{2BT} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left\{-\frac{(y(t_m)-s(t_m))^2}{2\sigma_n^2}\right\} \\ &= \frac{1}{(2\pi)^{BT}\sigma_n^{2BT}} \exp\left\{-\frac{1}{2\sigma_n^2}\sum_{m=1}^{2BT}(y(t_m)-s(t_m))^2\right\}. \end{aligned}$$

The log-likelihood ratio as

$$\begin{aligned} \ell(\underline{Y}) &= \log\left(\frac{f_1(y(t_1)\ldots y(t_{2BT}))}{f_0(y(t_1)\ldots y(t_{2BT}))}\right) \\ &= \log\left(\frac{\exp\{-\frac{1}{2\sigma_n^2}\sum_{m=1}^{2BT}(y(t_m)-s(t_m))^2\}}{\exp\{-\frac{1}{2\sigma_n^2}\sum_{m=1}^{2BT}y^2(t_m)\}}\right) \\ &= -\frac{1}{2\sigma_n^2}\left[\sum_{m=1}^{2BT}(y(t_m)-s(t_m))^2 - \sum_{m=1}^{2BT}y^2(t_m)\right]. \end{aligned}$$

Thus the most powerful test of size $\alpha = P_{\mathrm{FA}}$ is of the form

$$\frac{1}{2\sigma_w^2}\sum_{m=1}^{2BT}\left[2y(t_m)s(t_m) - s^2(t_m)\right] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_0,$$

where γ_0 is a threshold determined by the required false alarm rate P_{FA}.

Equivalently, we can write the test as

$$\frac{1}{N_0 B}\sum_{m=1}^{2BT} y(t_m)s(t_m) \underset{H_0}{\overset{H_1}{\gtrless}} \log\gamma_0 + \frac{1}{2N_0 B}\sum_{m=1}^{2BT} s^2(t_m).$$