ECE531 Lecture 5: Detection of Discrete-Time Signals with Unknown Parameters

D. Richard Brown III

Worcester Polytechnic Institute

23-February-2011

Introduction

- This lecture is about **parametric detection** problems for signals with one or more unknown parameters.
- We focus on the case of an N-sample discrete time observation $Y \in \mathcal{Y}$ with binary hypotheses

$$\begin{aligned} \mathcal{H}_0 &: Y \sim p_x(y) \text{ for } x \in \mathcal{X}_0 \\ \mathcal{H}_1 &: Y \sim p_x(y) \text{ for } x \in \mathcal{X}_1 \end{aligned}$$

- Our approach: Absorb the unknown parameters into our notion of the state x and the associated state space $\mathcal{X}$.
- In some textbooks, it is common to denote the random parameter(s) as Θ and a realization of these parameters as θ. We will stay consistent with our earlier notation, however, and use x here.

Example

Suppose we have a known signal $s \in \mathbb{R}^n$ and we have a communication system that transmits either $s_0 = -s$ or $s_1 = s$. The signal arrives with unknown amplitude $a > 0$ and is corrupted by zero-mean AWGN with variance σ^2. The hypotheses can be written as

$$\begin{aligned} \mathcal{H}_0 &: Y \sim \mathcal{N}(as_0, \sigma^2 I) \\ \mathcal{H}_1 &: Y \sim \mathcal{N}(as_1, \sigma^2 I) \end{aligned}$$

for unknown $a > 0$. What is the state space $\mathcal{X}$ and what are the sets $\mathcal{X}_0$ and $\mathcal{X}_1$?

An equivalent, but more clever way to write these hypotheses is

$$\begin{aligned} \mathcal{H}_0 &: Y \sim \mathcal{N}(as, \sigma^2 I) \text{ with } a < 0 \\ \mathcal{H}_1 &: Y \sim \mathcal{N}(as, \sigma^2 I) \text{ with } a > 0 \end{aligned}$$

for unknown $a \in \mathbb{R} \backslash 0$. What is the state space $\mathcal{X}$ and what are the sets $\mathcal{X}_0$ and $\mathcal{X}_1$?

Types of Detectors: What we Already Know

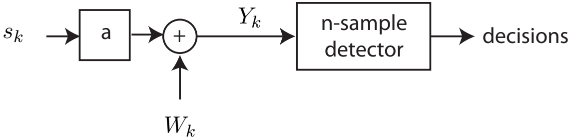
Problem: Find a decision rule to decide between

$$\begin{aligned}\mathcal{H}_0 &: Y \sim \mathcal{N}(as, \sigma^2 I) \text{ with } a < 0 \\ \mathcal{H}_1 &: Y \sim \mathcal{N}(as, \sigma^2 I) \text{ with } a > 0\end{aligned}$$

for unknown $a \in \mathbb{R} \backslash 0$

- We know that these types of problems are **composite** binary hypothesis testing problems.
- Bayes decision rules involve computation of two discriminant functions (or commodity costs) $g_0(y, \pi)$ and $g_1(y, \pi)$ and subsequent comparison.
- Under the N-P criterion with significance level (or size) α:
  - Our strategy: **reduce the composite HT problem to a simple one**.
  - Check for the existence of a UMP decision rule (check the critical region and/or monotone likelihood ratio) that maximizes $P_{D,x}$ for all $x \in \mathcal{X}_1$ subject to $P_{\text{fp},x} \leq \alpha$ for all $x \in \mathcal{X}_0$.
  - If a UMP rule doesn't exist, we can often find an LMP decision rule.

Example: Known Signal with Unknown Amplitude



- We would like to decide on the presence of a known signal $s \in \mathbb{R}^n$ ($\mathcal{H}_1$) versus the absence of the signal ($\mathcal{H}_0$).
- The known signal $s \in \mathbb{R}^n$ arrives at the detector corrupted by noise $W \in \mathbb{R}^n$ and with unknown amplitude $a \geq 0$.
- We observe a realization $y \in \mathbb{R}^n$ of the random variable $Y = as + W$.

$$\begin{aligned} a &= 0 \text{ when no signal is present } (\mathcal{H}_0) \\ a &> 0 \text{ when a signal is present } (\mathcal{H}_1) \end{aligned}$$

- What kind of hypothesis testing problem is this? Binary, composite, one-sided, with a simple null hypothesis.
- What is the state space here? $x = a \in \mathcal{X} = [0, \infty)$.

Bayes Detector for this Example

To develop a Bayes detector for this problem, we need 3 things:

1. We need a conditional pdf or pmf statistically describing the observations $y \in \mathcal{Y}$ for each state $x \in \mathcal{X}$: $p_x(y)$.
2. We need a cost assignment for each state $x \in \mathcal{X}$ and each hypothesis $\mathcal{H}_i$: $C_0(x)$ and $C_1(x)$.
3. We need a prior (pdf) on the states: $\pi(x)$.

The Bayes risk of decision rule ρ in this case can be written as

$$\begin{aligned} r(\rho, \pi) &= \int_{\mathcal{X}} \pi(x) \int_{\mathcal{Y}} \rho^\top(y) C(x) p_x(y)\, dy\, dx \\ &= \int_{\mathcal{Y}} \rho^\top(y) \underbrace{\left(\int_{\mathcal{X}} C(x)\pi(x) p_x(y)\, dx \right)}_{g(y,\pi)=[g_0(y,\pi), g_1(y,\pi)]^\top} dy \end{aligned}$$

We know how to solve this problem: For each $y \in \mathcal{Y}$, the Bayes decision rule just compares $g_0(y,\pi)$ to $g_1(y,\pi)$ and sets $\delta^{B\pi}$ equal to the index of the smaller discriminant function, i.e. commodity cost.

Bayes Detector for this Example

1. Suppose the noise is distributed as $W \sim \mathcal{N}(0, \Sigma)$. Conditioned on $x = a$, we can write the density of the vector observation as

$$p_{x=a}(y) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{(y-as)^\top \Sigma^{-1}(y-as)}{2}\right)$$

2. Suppose also that we have the UCA.
3. Suppose finally that we have the prior

$$\pi(x) = \pi_0\delta(x) + (1-\pi_0)\left(u(x) - u(x-1)\right)$$

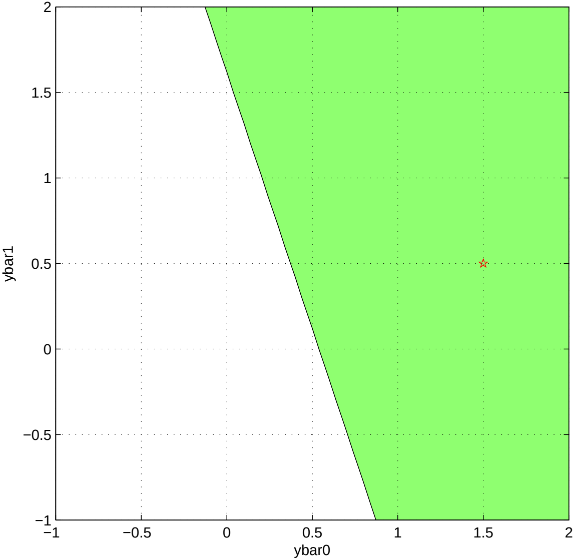
   where $u(x) = 1$ for all $x > 0$ and $u(x) = 0$ otherwise.

Note that the state space here is $\mathcal{X} = [0, 1]$.

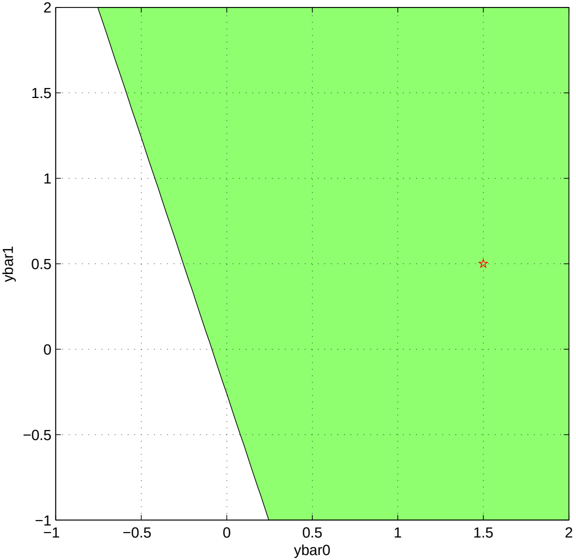
In this sort of problem, it is convenient to transform the observations (and decorrelate the noise) using our coordinate transformation trick.

```
% Bayes example plots
%---------------------------------------------------------------
% User variables
%---------------------------------------------------------------
pi0 = 0.75;            % prior state parameter
ybartest = -1:0.1:2;   % space of observations to test (in transformed coordinate space)
sbar = [1.5;0.5];      % signal vector (in transformed coordinate space)
%---------------------------------------------------------------

v = sbar'*sbar/2;
syms a real
discrimratio = zeros(length(ybartest),length(ybartest));

i0 = 0;
for y0 = ybartest
    i0 = i0+1;
    i1 = 0;
    for y1 = ybartest
    i1 = i1+1;
        ybar  = [y0;y1];
        u = sbar'*ybar;
        f=exp(a*u-a^2*v);
        discrimratio(i1,i0) = (1-pi0)/pi0 * int(f,0,1);
    end
end

contourf(ybartest,ybartest,discrimratio,[1 1]);
xlabel('ybar0'); ylabel('ybar1');
hold on; plot(sbar(1),sbar(2),'rp'); hold off; grid on; axis square
```

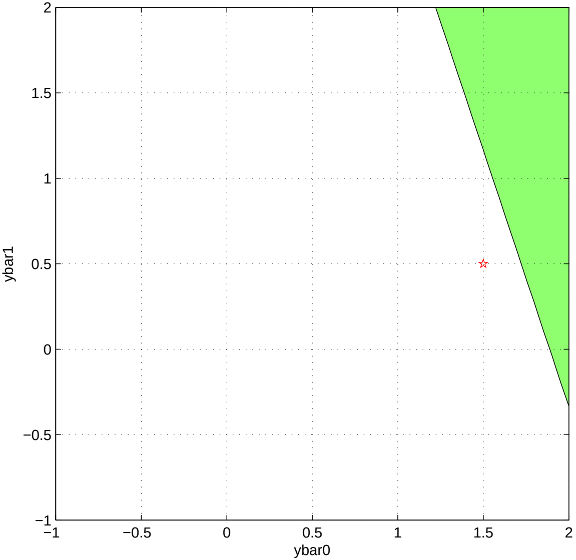
Bayes Decision Rule: $\pi_0 = 0.5$ and $\bar{s} = [1.5, 0.5]^\top$



Bayes Decision Rule: $\pi_0 = 0.4$ and $\bar{s} = [1.5, 0.5]^\top$



Bayes Decision Rule: $\pi_0 = 0.75$ and $\bar{s} = [1.5, 0.5]^\top$



Neyman-Pearson Detector

To develop a N-P detector for this problem, we need 2 things:

1. We need a conditional pdf or pmf statistically describing the observations $y \in \mathcal{Y}$ for each state $x \in \mathcal{X}$: $p_x(y)$.
2. We need a significance-level for the test: $P_{\mathsf{fp}} \leq \alpha$.

How should we approach the problem?

- Check to see if a UMP detector exists.
- If not, derive an LMP detector with good performance for a near zero.
- A new option ("hybrid"): If we can specify a prior on the states, denoted as $\pi(x)$, we can reduce each composite hypothesis to a simple one by taking a weighted average of the density with respect to the prior, i.e.

$$\mathcal{H}_j{}' \quad : \quad Y \sim p_j(y) = \int_{\mathcal{X}_j} p_x(y) \pi_j(x)\, dx$$

  where $\pi_j(x) = \frac{\pi(x)}{\mathrm{Prob}(x \in \mathcal{X}_j)}$ when $x \in \mathcal{X}_j$ and is equal to zero otherwise. Then testing $\mathcal{H}_0{}'$ versus $\mathcal{H}_1{}'$ is a simple hypothesis testing problem for which the N-P lemma gives a unique optimal solution.

N-P Detector: Known Signal with Unknown Amplitude

Suppose the noise is distributed as $W \sim \mathcal{N}(0, \sigma^2 I)$. Conditioned on $x = a$, we can write the density of the vector observation as

$$
\begin{aligned}
p_{x=a}(y) &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{(y-as)^\top (y-as)}{2\sigma^2}\right) \\
&= \prod_{k=0}^{n-1} p_{x=a}(y_k) \\
&= \prod_{k=0}^{n-1} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_k - as_k)^2}{2\sigma^2}\right)
\end{aligned}
$$

Note that any results we derive for this case can also be applied to the case when $W \sim \mathcal{N}(0, \Sigma)$ by using our coordinate-transformation (decorrelation) trick.

N-P Detector: Known Signal with Unknown Amplitude

To check for the existence of a UMP decision rule, let's set up a simple HT test: $\mathcal{H}_0 : a = 0$ versus $\mathcal{H}_1 : a = a_1 > 0$.

The N-P lemma says that, for this simple HT problem, we can find an α-level decision rule of the form

$$\rho(y) = \begin{cases} 1 & \ln\left(L(y)\right) > \ln v \\ \gamma & \ln\left(L(y)\right) = \ln v \\ 0 & \ln\left(L(y)\right) < \ln v \end{cases}$$

where $\ln\left(L(y)\right) := \ln\left(\frac{p_1(y)}{p_0(y)}\right)$ with $v \geq 0$ and $\gamma \in [0, 1]$ chosen such that $P_{\mathsf{fp}} = \alpha$.

Hence we need to compute the log-likelihood ratio...

N-P Detector: Known Signal with Unknown Amplitude

The log-likelihood ratio for this simple HT test:

$$
\begin{aligned}
\ln\left(L(y)\right) &= \ln\left(\frac{p_1(y)}{p_0(y)}\right) \\
&= \ln\left(\frac{\prod_{k=0}^{n-1}\exp\left(-\frac{(y_k-a_1 s_k)^2}{2\sigma^2}\right)}{\prod_{k=0}^{n-1}\exp\left(-\frac{y_k^2}{2\sigma^2}\right)}\right) \\
&= \ln\left(\prod_{k=0}^{n-1}\exp\left(\frac{2y_k a_1 s_k - a_1^2 s_k^2}{2\sigma^2}\right)\right) \\
&= \sum_{k=0}^{n-1}\frac{a_1}{\sigma^2}\left(y_k s_k - \frac{a_1 s_k^2}{2}\right) \\
&= \frac{a_1}{\sigma^2}\left(s^\top y - \frac{a_1||s||^2}{2}\right)
\end{aligned}
$$

Now we need to compute the threshold to achieve the false positive probability α.

N-P Detector: Known Signal with Unknown Amplitude

The false-positive probability can be calculated as

$$\begin{aligned} P_{\mathsf{fp}} &= \mathrm{Prob}\left(\ln\left(L(Y)\right) > \ln v\,;\, a=0\right) + \gamma \mathrm{Prob}\left(\ln\left(L(Y)\right) = \ln v\,;\, a=0\right) \\ &= \mathrm{Prob}\left(s^\top Y > \frac{\sigma^2}{a_1}\ln v + \frac{a_1\|s\|^2}{2}\,;\, a=0\right) \end{aligned}$$

How is $s^\top Y$ distributed when $a=0$?

Hence, v must be selected such that

$$P_{\mathsf{fp}} = Q\left(\frac{\frac{\sigma^2}{a_1}\ln v + \frac{a_1\|s\|^2}{2}}{\|s\|\sigma}\right) = \alpha$$

This implies the a value for v.

N-P Detector: Known Signal with Unknown Amplitude

Does a UMP decision rule exist? To answer this question, we need to determine if the critical region $\Gamma_1 = \{y \in \mathcal{Y} : \rho(y) \text{ decides } \mathcal{H}_1\}$ depends on our choice of a_1.

We decide $\mathcal{H}_1$ for sure when

$$s^\top y > \frac{\sigma^2}{a_1} \ln v + \frac{a_1 ||s||^2}{2}$$

and we decide $\mathcal{H}_1$ with probability γ when

$$s^\top y = \frac{\sigma^2}{a_1} \ln v + \frac{a_1 ||s||^2}{2}$$

(which happens with probability zero).

Critical region depends on $a_1 \Rightarrow$ no UMP decision rule. ☹
What should we do now?

N-P Detector: Known Signal with Unknown Amplitude

Hold on. Does the critical region $\Gamma_1 = \{y \in \mathbb{R}^n : s^\top y > \frac{\sigma^2}{a_1} \ln v + \frac{a_1 ||s||^2}{2}\}$ really depend on a_1?

Given $0 \le \alpha \le 1$ and $t > 0$, the unique solution to $Q(z/t) = \alpha$ is $z = tQ^{-1}(\alpha)$. Hence, the unique solution to

$$Q\left(\frac{\frac{\sigma^2}{a_1}\ln v + \frac{a_1||s||^2}{2}}{||s||\sigma}\right) = \alpha$$

is

$$\frac{\sigma^2}{a_1}\ln v + \frac{a_1||s||^2}{2} = ||s||\sigma Q^{-1}(\alpha).$$

Hence the critical region for a significance-level α N-P decision rule can be written as

$$\Gamma_1 = \{y \in \mathbb{R}^n : s^\top y > ||s||\sigma Q^{-1}(\alpha)\}$$

Does this depend on a_1? Does a UMP decision rule exist?

N-P Detector: Known Signal with Unknown Amplitude

Here is our UMP decision rule:

$$\rho^{\mathsf{UMP}}(y) = \begin{cases} 1 & s^\top y > ||s||\sigma Q^{-1}(\alpha) \\ \gamma & s^\top y = ||s||\sigma Q^{-1}(\alpha) \\ 0 & s^\top y < ||s||\sigma Q^{-1}(\alpha) \end{cases}$$

How would this change if the noise was distributed as $W \sim \mathcal{N}(0, \Sigma)$?

N-P Detector: Known Signal with Unknown Amplitude

If the UMP detector did not exist or was too complicated, we could find an LMP detector for this example by comparing

$$\frac{d}{da} L_a(y)|_{a=0}$$

to a threshold. When the noise samples are i.i.d., the likelihood ratio can be written as

$$\begin{aligned} L_a(y) &= \frac{p_{x=a}(y)}{p_{x=0}(y)} \\ &= \prod_{k=0}^{n-1} \frac{q(y_k - a s_k)}{q(y_k)} \end{aligned}$$

where $q(x)$ is the pmf/pdf of the kth noise sample. In our example, $q(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/\sigma^2}$. We'll continue our analysis here for i.i.d. noise with general distribution $q(x)$...

Neyman-Pearson Detector: LMP Decision Rule

Taking the derivative of $L_a(y)$ with respect to a yields

$$\frac{d}{da}L_a(y) = -\sum_{k=0}^{n-1} s_k \frac{q'(y_k - as_k)}{q(y_k)} \prod_{j\neq k} \frac{q(y_j - as_j)}{q(y_j)}$$

where $q'(x) = \frac{d}{dx}q(x)$. Setting $a = 0$ yields

$$\begin{aligned} \frac{d}{da}L_a(y)\Big|_{a=0} &= -\sum_{k=0}^{n-1} s_k \frac{q'(y_k)}{q(y_k)} \prod_{j\neq k} \frac{q(y_j)}{q(y_j)} \\ &= -\sum_{k=0}^{n-1} s_k \frac{q'(y_k)}{q(y_k)} \\ &= \sum_{k=0}^{n-1} s_k h^{\mathsf{LMP}}(y_k) \end{aligned}$$

Neyman-Pearson Detector: LMP Decision Rule

The locally most powerful decision rule then takes the form

$$\rho^{\mathsf{LMP}}(y) = \begin{cases} 1 & \sum_{k=0}^{n-1} s_k h^{\mathsf{LMP}}(y_k) > \tau \\ \gamma & \sum_{k=0}^{n-1} s_k h^{\mathsf{LMP}}(y_k) = \tau \\ 0 & \sum_{k=0}^{n-1} s_k h^{\mathsf{LMP}}(y_k) < \tau \end{cases}$$

where γ and τ are selected such that $P_{\mathsf{fp}} = \alpha$.

In our Gaussian example, $h^{\mathsf{LMP}}(y_k) := -\frac{q'(y_k)}{q(y_k)} = \frac{2y_k}{\sigma^2}$. Hence, if $\tau' = \tau\sigma^2/2$, the LMP decision rule then takes the form

$$\rho^{\mathsf{LMP}}(y) = \begin{cases} 1 & s^\top y > \tau' \\ \gamma & s^\top y = \tau' \\ 0 & s^\top y < \tau' \end{cases}$$

with τ' and γ selected so that $P_{\mathsf{fp}} = \alpha$. How does this compare to the UMP decision rule? $\rho^{\mathsf{LMP}}(y) = \rho^{\mathsf{UMP}}(y)$ when the noise is i.i.d. Gaussian.

LMP Decision Rule for i.i.d. Laplacian Noise

For $b > 0$, the Laplacian density is given as $q(x) = \frac{b}{2}e^{-b|x|}$.
For Laplacian noise,

$$h^{\mathsf{LMP}}(y_k) = -\frac{q'(y_k)}{q(y_k)} = b\,\mathsf{sgn}(x).$$

Hence, the locally most powerful decision rule can be written as

$$\rho^{\mathsf{LMP}}(y) = \begin{cases} 1 & \sum_{k=0}^{n-1} s_k \mathsf{sgn}(y_k) > \tau \\ \gamma & \sum_{k=0}^{n-1} s_k \mathsf{sgn}(y_k) = \tau \\ 0 & \sum_{k=0}^{n-1} s_k \mathsf{sgn}(y_k) < \tau \end{cases}$$

In this case, $\rho^{\mathsf{LMP}}(y) \neq \rho^{\mathsf{UMP}}(y)$.

Hybrid Approach: Average Density Over a Prior

An sub-optimal approach used in some textbooks is to reduce the composite HT problem to a simple HT problem by using the a prior $\pi(x)$ to compute a single density as a weighted average of the family of densities associated with $\mathcal{H}_j$, i.e.

$$p_j(y) = \int_{\mathcal{X}_j} p_x(y)\pi_j(x)\,dx$$

where $\pi_j(x) = \frac{\pi(x)}{\mathrm{Prob}(x \in \mathcal{X}_j)}$ when $x \in \mathcal{X}_j$ and is equal to zero otherwise. In this case, given a significance level α, the N-P lemma tells us that a unique optimal solution must exist based on a threshold test of the likelihood ratio

$$\begin{aligned} L(y) &= \frac{p_1(y)}{p_0(y)} \\ &= \frac{\int_{\mathcal{X}_1} p_x(y)\pi_1(x)\,dx}{\int_{\mathcal{X}_0} p_x(y)\pi_0(x)\,dx}. \end{aligned}$$

Remarks on the Hybrid Approach

- If there is some uncertainty as to the prior, the usual approach is to chose $\pi(x)$ such that it gives as little information about x as possible, i.e. such that the simple HT problem has "maximum difficulty".
- Suppose the unknown parameter is the phase of the received signal. What prior on this phase would give the least information?

$$\pi(x) \sim \mathcal{U}(0, 2\pi) \text{ or } \pi(x) \sim \mathcal{U}(-\pi, \pi)$$

This is commonly used in practice.

Conclusions

- Detection of known signals in noise: **simple hypothesis testing**.
- Detection of signals with one or more unknown parameters in noise: **composite hypothesis testing**.
- Your textbook has lots of examples in Chapter 7. Unfortunately, most of them are solved with the GLRT.
- The GLRT is not covered in ECE531, but you should know that it is a suboptimal technique in which
  - The unknown parameters are first **estimated** (to be covered in the second half of this course)
  - The estimates are plugged in to the signal model so now all of the parameters are "known"
  - Simple hypothesis testing is performed

Midterm Exam: What You Need to Know

- A basic understanding of the different types of hypothesis testing problems (binary, M-ary, simple, composite).
- Mathematical model of hypothesis testing problems.
- Intuition about "good" and "bad" decision rules.
- Lecture notes from lectures 1-4 (not including LMP detectors)
  - Bayesian hypothesis testing (binary and M-ary, simple and composite)
  - Neyman-Pearson hypothesis testing (binary, simple and composite)
  - Decorrelation of signals observed in correlated Gaussian noise.
  - Detection of known discrete-time signals (binary and M-ary)
  - Detection of discrete-time signals with random parameters (binary)
- Kay II Chapters 1, 3, 4, 6 but not including
  - GLRT, Rao, Wald detectors for composite HT problems
  - LMP decision rules
  - Anything having to do with estimation