ELE530: Estimation

Fernando Pérez-Cruz

Princeton University and University Carlos III in Madrid

March 25th, 2009
From Detection to Estimation

- We want to know if a coin bias is $P(H) = 0.2$ or $P(H) = 0.3$ or $P(H) = 0.35$.

- We want to know if the bit $s$ is zero or one:

$$x = y \cdot s + w$$

- We want to know if a target is present.
From Detection to Estimation

- We want to know if a coin bias is $P(H) = 0.2$ or $P(H) = 0.3$ or $P(H) = 0.35$.
  - We want to know the coin bias, $P(H)$.
- We want to know if the bit $s$ is zero or one:
  \[ x = y \cdot s + w \]
- We want to know if a target is present.
From Detection to Estimation

- We want to know if a coin bias is $P(H) = 0.2$ or $P(H) = 0.3$ or $P(H) = 0.35$.
  - We want to know the coin bias, $P(H)$.

- We want to know if the bit $s$ is zero or one:
  \[ x = y \cdot s + w \]
  - We want to know the gain of the channel $y$.

- We want to know if a target is present.
From Detection to Estimation

- We want to know if a coin bias is $P(H) = 0.2$ or $P(H) = 0.3$ or $P(H) = 0.35$.
  - We want to know the coin bias, $P(H)$.

- We want to know if the bit $s$ is zero or one:

  $$x = y \cdot s + w$$
  
  - We want to know the gain of the channel $y$.

- We want to know if a target is present.
  
  - We want to know the position and velocity of the target.
Minimum Risk

Detection:

\[ R(\delta) = \sum_y \int L(\delta(x), y)p(x, y)dx \]

where \( L(\delta(x), y) \) is a \( M \times M \) matrix, measuring the penalty from \( \delta(x) \) to \( y \).

Estimation:

\[ R(\delta) = \int L(\delta(x), y)p(x, y)dxdy \]

where \( L(\delta(x), y) \) is function \( \mathbb{R} \times \mathbb{R} \to \mathbb{R} \), measuring the penalty from \( \delta(x) \) to \( y \).
Loss Functions

- **Minimum Error:**
  - $L(\delta(x), y) = L(\delta(x) - y)$.
  - $L(y - \delta(x)) = L(\delta(x) - y)$.
  - $L(e_1) \leq L(e_2)$ for $|e_1| < |e_2|$.
Loss Functions

- Minimum Error:
  - \( L(\delta(x), y) = L(\delta(x) - y) \).
  - \( L(y - \delta(x)) = L(\delta(x) - y) \).
  - \( L(e_1) \leq L(e_2) \) for \( |e_1| < |e_2| \).

- Standard Loss functions:
  - Squared Error: \( L(\delta(x), y) = (\delta(x) - y)^2 \).
  - Absolute Error: \( L(\delta(x), y) = |\delta(x) - y| \).
  - Uniform Error: \( L(\delta(x), y) = \begin{cases} 0, & |\delta(x) - y| \leq \Delta/2 \\ 1, & |\delta(x) - y| > \Delta/2 \end{cases} \).
Minimum Risk Estimation Rule

- Minimum Risk:
  \[
  \delta_{\text{opt}} = \arg\min_{\delta} \int L(\delta(x), y)p(x, y)dx dy
  \]
  \[
  = \arg\min_{\delta} \int p(x)dx \int L(\delta(x), y)p(y|x)dy
  \]

- We can get the minimum risk rule for each \(x\):
  \[
  \delta_{\text{opt}}(x) = \arg\min_{\delta(x)} \int L(\delta(x), y)p(y|x)dy
  \]

- This solution is identical to the detection problem. We define a minimum risk estimation rule for each \(x\).
Minimum Mean Squared Error

- Squared error $L(\delta(x) - y) = (\delta(x) - y)^2$:

$$\hat{\delta}_{\text{mmse}}(x) = \arg \min_{\delta(x)} \int (\delta(x) - y)^2 p(y|x) dy$$
Minimum Mean Squared Error

- Squared error \( L(\delta(x) - y) = (\delta(x) - y)^2 \):

\[
\delta_{\text{mmse}}(x) = \arg \min_{\delta(x)} \int (\delta(x) - y)^2 p(y|x) dy
\]

\[
= \arg \min_{\delta(x)} \delta^2(x) - 2\delta(x) \int y p(y|x) dy + \int y^2 p(y|x) dy
\]
Minimum Mean Squared Error

Squared error \( L(\delta(x) - y) = (\delta(x) - y)^2 \):

\[
\delta_{mmse}(x) = \arg \min_{\delta(x)} \int (\delta(x) - y)^2 p(y|x) dy \\
= \arg \min_{\delta(x)} \delta^2(x) - 2\delta(x) \int y p(y|x) dy + \int y^2 p(y|x) dy \\
= \arg \min_{\delta(x)} \delta^2(x) - 2\delta(x) E[y|x] + \int y^2 p(y|x) dy
\]
Minimum Mean Squared Error

Square error $L(\delta(x) - y) = (\delta(x) - y)^2$:

$$\delta^{\text{mmse}}(x) = \arg\min_{\delta(x)} \int (\delta(x) - y)^2 p(y|x) dy$$

$$= \arg\min_{\delta(x)} \delta^2(x) - 2\delta(x) \int yp(y|x) dy + \int y^2 p(y|x) dy$$

$$= \arg\min_{\delta(x)} \delta^2(x) - 2\delta(x) E[y|x] + \int y^2 p(y|x) dy$$

Now we take the derivative with respect to $\delta(x)$:

$$2\delta^{\text{mmse}}(x) - 2E[y|x] = 0$$
Minimum Mean Squared Error

- Squared error $L(\delta(x) - y) = (\delta(x) - y)^2$:

$$\delta_{\text{mmse}}(x) = \arg \min_{\delta(x)} \int (\delta(x) - y)^2 p(y|x) dy$$

$$= \arg \min_{\delta(x)} \delta^2(x) - 2\delta(x) \int yp(y|x) dy + \int y^2 p(y|x) dy$$

$$= \arg \min_{\delta(x)} \delta^2(x) - 2\delta(x) E[y|x] + \int y^2 p(y|x) dy$$

- Now we take the derivative with respect to $\delta(x)$:

$$2\delta_{\text{mmse}}(x) - 2E[y|x] = 0$$

$$\delta_{\text{mmse}}(x) = E[y|x]$$

- The MMSE solution is the Conditional Mean Estimate.
Minimum Mean Absolute Error

Absolute error \( L(\delta(x) - y) = |\delta(x) - y| \):

\[
\delta_{\text{mmae}}(x) = \arg \min_{\delta(x)} \int |\delta(x) - y| p(y|x) dy
\]
Minimum Mean Absolute Error

- Absolute error \( L(\delta(x) - y) = |\delta(x) - y| : \)

\[
\delta^{\text{mmae}}(x) = \arg \min_{\delta(x)} \int |\delta(x) - y| p(y|x) dy
\]

\[
= \arg \min_{\delta(x)} \left( \int_{-\infty}^{\delta(x)} (\delta(x) - y) p(y|x) dy - \int_{\delta(x)}^{\infty} (\delta(x) - y) p(y|x) dy \right)
\]
Minimum Mean Absolute Error

Absolute error \( L(\delta(x) - y) = |\delta(x) - y| : \)

\[
\delta_{\text{mmae}}(x) = \arg \min_{\delta(x)} \int |\delta(x) - y| p(y|x) dy \\
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} (\delta(x) - y)p(y|x) dy - \int_{\delta(x)}^{\infty} (\delta(x) - y)p(y|x) dy \\
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} P(Y \leq y|x) dy - \int_{\delta(x)}^{\infty} P(Y \leq y|x) dy
\]
Minimum Mean Absolute Error

- Absolute error $L(\delta(x) - y) = |\delta(x) - y|$: 

$$
\delta_{\text{mmae}}(x) = \arg \min_{\delta(x)} \int |\delta(x) - y| p(y|x) dy \\
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} (\delta(x) - y) p(y|x) dy - \int_{\delta(x)}^{\infty} (\delta(x) - y) p(y|x) dy \\
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} p(Y \leq y|x) dy - \int_{\delta(x)}^{\infty} p(Y \leq y|x) dy
$$

- Now we take the derivative with respect to $\delta(x)$: 

$$
P(Y \leq \delta_{\text{mmae}}(x)|x) - P(Y > \delta_{\text{mmae}}(x)|x) = 0
$$
Minimum Mean Absolute Error

- Absolute error $L(\delta(x) - y) = |\delta(x) - y|$: 

$$
\delta_{\text{mmae}}(x) = \arg \min_{\delta(x)} \int |\delta(x) - y| p(y|x) dy
$$

$$
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} (\delta(x) - y) p(y|x) dy - \int_{\delta(x)}^{\infty} (\delta(x) - y) p(y|x) dy
$$

$$
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)} P(Y \leq y|x) dy - \int_{\delta(x)}^{\infty} P(Y \leq y|x) dy
$$

- Now we take the derivative with respect to $\delta(x)$:

$$
P(Y \leq \delta_{\text{mmae}}(x)|x) = 1/2
$$

- The MMAE solution is the Conditional Median Estimate.
Minimum Mean Uniform Error

Uniform error \( L(\delta(x), y) = \begin{cases} 0, & |\delta(x) - y| \leq \Delta/2, \\ 1, & |\delta(x) - y| > \Delta/2 \end{cases} \)

\[ \delta_{\text{mmue}}(x) = \arg \min_{\delta(x)} \int L(\delta(x) - y)p(y|x)dy \]

\[ = \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x) - \Delta/2} p(y|x)dy + \int_{\delta(x) + \Delta/2}^{\infty} p(y|x)dy \]
Minimum Mean Uniform Error

- Uniform error \( L(\delta(x), y) = \begin{cases} 
0, & |\delta(x) - y| \leq \Delta/2 \\
1, & |\delta(x) - y| > \Delta/2 
\end{cases} \)

\[
\delta_{mmue}(x) = \arg\min_{\delta(x)} \int L(\delta(x) - y)p(y|x)dy
\]

\[
= \arg\min_{\delta(x)} \int_{\delta(x) - \Delta/2}^{\infty} p(y|x)dy + \int_{\delta(x) + \Delta/2}^{\infty} p(y|x)dy
\]

\[
= \arg\min_{\delta(x)} 1 - [P(Y \leq \delta(x) + \Delta/2|x) - P(Y \leq \delta(x) - \Delta/2|x)]
\]
Minimum Mean Uniform Error

Uniform error \( L(\delta(x), y) = \begin{cases} 0, & |\delta(x) - y| \leq \Delta/2 \\ 1, & |\delta(x) - y| > \Delta/2 \end{cases} \)

\[
\delta_{mmue}(x) = \arg \min_{\delta(x)} \int L(\delta(x) - y) p(y|x) dy
\]

\[
= \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x) - \Delta/2} p(y|x) dy + \int_{\delta(x) + \Delta/2}^{\infty} p(y|x) dy
\]

\[
= \arg \min_{\delta(x)} 1 - [P(Y \leq \delta(x) + \Delta/2|x) - P(Y \leq \delta(x) - \Delta/2|x)]
\]

We can instead maximize:

\[
\delta_{mmue}(x) = \arg \max_{\delta(x)} \frac{P(Y \leq \delta(x) + \Delta/2|x) - P(Y \leq \delta(x) - \Delta/2|x)}{\Delta}
\]
Minimum Mean Uniform Error

Uniform error \( L(\delta(x), y) = \begin{cases} 
0, & |\delta(x) - y| \leq \Delta/2 \\
1, & |\delta(x) - y| > \Delta/2 
\end{cases} \)

\[
\delta_{\text{mmue}}(x) = \arg\min_{\delta(x)} \int L(\delta(x) - y)p(y|x)dy
\]

\[
= \arg\min_{\delta(x)} \int_{-\infty}^{\delta(x) - \Delta/2} p(y|x)dy + \int_{\delta(x)+\Delta/2}^{\infty} p(y|x)dy
\]

\[
= \arg\min_{\delta(x)} 1 - \left[ P(Y \leq \delta(x) + \Delta/2|x) - P(Y \leq \delta(x) - \Delta/2|x) \right]
\]

We can instead maximize:

\[
\delta_{\text{mmue}}(x) = \arg\max_{\delta(x)} \frac{P(Y \leq \delta(x) + \Delta/2|x) - P(Y \leq \delta(x) - \Delta/2|x)}{\Delta}
\]

\[
= \arg\max_{y} p(y|x)
\]
Minimum Mean Uniform Error

- Uniform error:
  \[ \delta_{\text{mmue}}(x) = \arg \min_{\delta(x)} \int_{-\infty}^{\delta(x)-\Delta/2} p(y|x)dy + \int_{\delta(x)+\Delta/2}^{\infty} p(y|x)dy \]

- We can instead maximize:
  \[ \delta_{\text{mmue}}(x) = \arg \max_y p(y|x) \]

- Now we take the derivative with respect to \( \delta(x) \):
  \[ \frac{dp(y|x)}{dy} \bigg|_{y=\delta_{\text{mmue}}(x)} = 0 \]

- The MMUE solution is the Conditional Mode Estimate.

- The MMUE solution is the Maximum a posteriori (MAP).
Loss Functions
Example

- Estimate the MMSE, MMAE and MMUE rules:

\[ p(y) = \alpha e^{-\alpha y} \quad y \geq 0 \]
\[ p(x|y) = ye^{-yx} \quad x \geq 0 \]

- First, we compute the joint density:

\[ p(x, y) = \alpha ye^{-(\alpha + x)y} \quad x, y \geq 0 \]

- Second, we derive the conditional density for \( y \):

\[ p(y|x) = \frac{p(y, x)}{p(x)} = (\alpha + x)^2 ye^{-(\alpha + x)y} \quad y \geq 0 \]
Example: Squared Error

Estimate the minimum mean squared error rule:

\[ \delta_{\text{mmse}}(x) = \int_0^{\infty} y^2(\alpha + x)^2 e^{-(\alpha+x)y} dy \]

\[ = -y^2(\alpha + x)e^{-(\alpha+x)y} \bigg|_0^{\infty} + 2 \int_0^{\infty} y(\alpha + x)e^{-(\alpha+x)y} dy \]

\[ = -2ye^{-(\alpha+x)y} \bigg|_0^{\infty} + 2 \int_0^{\infty} e^{-(\alpha+x)y} dy \]

\[ = -\frac{2}{\alpha + x} e^{-(\alpha+x)y} \bigg|_0^{\infty} \]

\[ = \frac{2}{\alpha + x} \]
Example: Absolute Error

Estimate the minimum mean absolute error rule:

\[ \int_{\delta_{\text{mmae}}(x)}^{\infty} y(\alpha + x)^2 e^{-(\alpha+x)y} dy = \frac{1}{2} \]

\[ -y(\alpha + x)e^{-(\alpha+x)y}\bigg|_{\delta_{\text{mmae}}(x)}^{\infty} + \int_{\delta_{\text{mmae}}(x)}^{\infty} (\alpha + x)e^{-(\alpha+x)y} dy = \frac{1}{2} \]

\[ \delta_{\text{mmae}}(x)(\alpha + x)e^{-(\alpha+x)\delta_{\text{mmae}}(x)} + e^{-(\alpha+x)\delta_{\text{mmae}}(x)} = \frac{1}{2} \]

\[ \delta_{\text{mmae}}(x) = \frac{K}{\alpha + x} \quad \text{where} \quad (K + 1)e^{-K} = \frac{1}{2} \]

\[ \delta_{\text{mmae}}(x) = \frac{1.6783}{\alpha + x} \]
Example: Uniform Error

Estimate the minimum mean uniform error rule:

\[
\left. \frac{dp(y|x)}{dy} \right|_{y=\delta_{mmue}(x)} = (\alpha + x)^2 e^{-(\alpha + x)y(1 - y(\alpha + x))} = 0
\]

\[
1 - \delta_{mmae}(x)(\alpha + x) = 0
\]

\[
\delta_{mmae}(x) = \frac{1}{\alpha + x}
\]

Estimators:

<table>
<thead>
<tr>
<th>MMSE</th>
<th>MMAE</th>
<th>MMUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{2}{\alpha + x} )</td>
<td>( \frac{1.6783}{\alpha + x} )</td>
<td>( \frac{1}{\alpha + x} )</td>
</tr>
</tbody>
</table>
Example 2

- Estimate the MMSE, MMAE and MMUE rules:

\[
    p(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1 - y)^{\beta-1} \quad 0 \leq y \leq 1
\]

\[
    p(x|y) = ye^{-yx} \quad x \geq 0
\]

- First, we compute the joint density:

\[
    p(x, y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha}(1 - y)^{\beta-1} e^{-xy} \quad x \geq 0, \ 0 \leq y \leq 1
\]

- To derive the conditional density for \( y \), we need

\[
    p(x) = \int_{0}^{1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha}(1 - y)^{\beta-1} e^{-xy} dy
\]
Example 2: Uniform Error

Estimate the minimum mean uniform error rule:

\[
\frac{dp(y|x)}{dy} \bigg|_{y=\delta_{mmue}(x)} = \frac{dp(y,x)}{p(x)} \bigg|_{y=\delta_{mmue}(x)} = \frac{dp(y,x)}{dy} \bigg|_{y=\delta_{mmue}(x)}
\]

\[
= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}y^{\alpha-1}(1 - y)^{\beta-2}e^{-xy}\left(\alpha(1 - y) - (\beta - 1)y - xy(1 - y)\right) = 0
\]

\[
\alpha(1 - y) - (\beta - 1)y - xy(1 - y) = 0
\]

\[
xy^2 - (x + \alpha + \beta - 1)y + \alpha = 0
\]

\[
\delta_{mmue}(x) = \frac{(x + \alpha + \beta - 1) - \sqrt{(x + \alpha + \beta - 1)^2 - 4\alpha x}}{2x}
\]
Properties

Property 1:
If \( L(\delta(x) - y) \) is symmetric and convex, and \( p(y|x) \) is symmetric about its conditional mean, then
\[
\delta(x) = \arg \min_{\delta(x)} \int L(\delta(x) - y)p(y|x)dy = \delta_{\text{mmse}}(x)
\]

Property 2:
If \( L(\delta(x) - y) \) is symmetric and nondecreasing, \( p(y|x) \) is symmetric about its conditional mean and unimodal, and \( \lim_{a \to \infty} L(a)p(a|x) = 0 \), then
\[
\delta(x) = \arg \min_{\delta(x)} \int L(\delta(x) - y)p(y|x)dy = \delta_{\text{mmse}}(x)
\]
Nonrandom Parameter Estimation

- Now assume $y$ is an unknown parameter.
- Therefore, there is not a $p(y)$ that characterizes $y$.
- What can we do?
Nonrandom Parameter Estimation

► Now assume $y$ is an unknown parameter.

► Therefore, there is not a $p(y)$ that characterizes $y$.

► What can we do?
  
  • Is there an equivalent Neyman-Pearson rule for estimation?
Nonrandom Parameter Estimation

Now assume $y$ is an unknown parameter.

Therefore, there is not a $p(y)$ that characterizes $y$.

What can we do?

- Is there an equivalent Neyman-Pearson rule for estimation?

- Is there an equivalent Minimax rule for estimation?
Nonrandom Parameter Estimation

Now assume $y$ is an unknown parameter.

Therefore, there is not a $p(y)$ that characterizes $y$.

What can we do?

- Is there an equivalent Neyman-Pearson rule for estimation?
  
  No

- Is there an equivalent Minimax rule for estimation?
  
  Somewhat.
Nonrandom Parameter Estimation

Now assume $y$ is an unknown parameter.

Therefore, there is not a $p(y)$ that characterizes $y$.

What can we do?

- Is there an equivalent Neyman-Pearson rule for estimation?
  No, because it does not make sense to impose $P_{fa} \leq \alpha$, there is no $P_{fa}$ or $P_M$ for estimation problems.

- Is there an equivalent Minimax rule for estimation?
  Somewhat.
Nonrandom Parameter Estimation

Now assume \( y \) is an unknown parameter.

Therefore, there is not a \( p(y) \) that characterizes \( y \).

What can we do?

- Is there an equivalent Neyman-Pearson rule for estimation?
  No, because it does it not make sense to impose \( P_{fa} \leq \alpha \), there is no \( P_{fa} \) or \( P_M \) for estimation problems.

- Is there an equivalent Minimax rule for estimation?
  Somewhat. For parameters with finite support we can define \( p(y) = d(y-y_0) \) and look for the estimation rule that minimizes the maximum error. Does it makes sense?
Difference Between Classic and Bayesian Statistics