Lecture 15 - Outline

1. Decision theory
   a. Binary hypothesis testing
   b. Maximum Likelihood decisions for discrete RVs
   c. Maximum Likelihood decisions for continuous RVs
   d. Bayesian hypothesis testing

Motivating Example.  Radar/Sonar Detection

- A radar/sonar system transmits electromagnetic/ultrasonic pulses in the direction of a target, which may (or may not) be "out there".
- The receiver listens for echoes
  - Echoes may result from the target or from other extraneous effects (noise, large birds, atmospheric inhomogeneities, large metallic flying objects, ...)
- Based on observed echoes, the radar system decides whether a target is present or not

Simple Concepts

- System decides if a target is present or not
- The decision may be correct or incorrect
- If the decision is that a target is present
  - This decision may be wrong if there is no target present
    - FALSE ALARM
- If decision is that no target is present
  - This decision may be wrong if there is a target present
    - MISSED DETECTION

8.1 Binary Test of Hypothesis

Decide Between Two Hypotheses $H_0$ and $H_1$, e.g., Radar: $H_0$: a target is absent, and $H_1$: a target is present

One and only one of these hypotheses is true, i.e., $H_0$ and $H_1$ are mutually exclusive exhaustive...

Decision is based on an observation, a random variable $X$ (discrete or continuous): $X$ is statistically dependent on which hypothesis is true.

The conditional probabilities $P[X|H_0]$ and $P[X|H_1]$ are known & are called the likelihoods.

Decision

Decision maker selects one of the hypotheses. The event that it selects $H_0$ is called $U_0$. The event that it selects $H_1$ is called $U_1$.

Decision rule maps $X$ into decision $U$: $U \in \{U_0, U_1\}$

- $U$ is derived random variable!
Decision Errors
1) System selects $H_1$ when $H_0$ is true (target is present when it is actually absent): FALSE ALARM
2) System selects $H_0$ when $H_1$ is true (target is absent when it is actually present): MISSED DETECTION

Average Probability of Error: Use Total Probability Theorem
$$P[\text{Error}] = P[\text{Error}|H_1]P[H_1] + P[\text{Error}|H_0]P[H_0]$$

Equal a priori probabilities on $H_0, H_1$: $P[H_0] = P[H_1]$

Goal: Develop a decision strategy that reduces these errors

Example 2. Significance Testing

- Consider a system (e.g., a production process) with a random output (e.g., defective items)
- Want to characterize the output using a probability model
- Suppose we postulate (or estimate from past measurements) the output to follow some probability model (e.g., defectives follow Bernoulli w.p. $p_0=0.01$)
- Introduce a change in the system. Decide:
  $H_0$ (null hypothesis): $p=p_0$ or $H_1$: $p\neq p_0$
- Significance Level $\alpha \leq$ False alarm probability $P[U_1|H_0]$

Rationale for the ML Decision Rule
- Select the hypothesis that maximizes the probability of obtaining the observed value of $X$
- It inherently assumes that the hypotheses are a priori equally likely

An Equivalent Implementation of ML Decision Rule

Likelihood Ratio $L(k) = \frac{P[X=x_k|H_1]}{P[X=x_k|H_0]}$

Chose $H_1$ if $L(k) > 1$
Chose $H_0$ if $L(k) < 1$

Example 1. Binary Channels

- Modern transmits bits over a noisy channel
  - Bits are 0 (hypothesis $H_0$) or 1 (hypothesis $H_1$)
- Output of channel: random variable $X$ generated for each bit, independently, conditioned on value of input bit
- Probability law of $X$ depends on input to channel (0 or 1)
- Decision problem: decode each bit given observation $X$

8.2 Maximum-Likelihood Decision for Discrete RVs

Observation RV $X$ takes on values $x_1, x_2, \ldots x_N$

Write the PMFs for each of the two hypotheses in an array, called the likelihood array, or likelihood matrix:

<table>
<thead>
<tr>
<th>$X = x_1$</th>
<th>$X = x_2$</th>
<th>$X = x_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P[X</td>
<td>H_0]$</td>
<td>$P[X</td>
</tr>
</tbody>
</table>

It is an array of 2 rows and $N$ columns, one row per hypothesis and one column per measurement value

Maximum-Likelihood Decision Rule:
If the observation is $X = x_k$, then
- Chose $H_k$ if $P[X=x_k|H_k] > P[X=x_k|H_l]$ for all $l \neq k$
- (If tie, select either)

Yet Another Equivalent Implementation:

Log-Likelihood Ratio
$$\log L(k) = \log P[X=x_k|H_1] - \log P[X=x_k|H_0]$$

Chose $H_1$ if $\log L(k) > 0$
Chose $H_0$ if $\log L(k) < 0$
Hypotheses:

- \( H_0 \): coin is biased with probability of heads = 0.2
- \( H_1 \): coin is fair (probability of heads = 0.5)

Observation: \( X \) = number of heads in 3 flips

Likelihoods:
- \( P[X|H_0] = \text{Binom}(3, 0.2) \)
- \( P[X|H_1] = \text{Binom}(3, 0.5) \)

Likelihood matrix:

\[
\begin{array}{cccc}
X = 0 & X = 1 & X = 2 & X = 3 \\
P[X|H_0] & 0.125 & 0.375 & 0.375 & 0.125 \\
P[X|H_1] & 0.125 & 0.375 & 0.375 & 0.125 \\
\end{array}
\]

ML Decision Rule: Find and shade the largest entry in each column!  

Error Probabilities:
- \( P[\text{error}|H_0] = P[X=2|H_0] + P[X=3|H_0] = 0.096 + 0.008 = 0.104 \)
- \( P[\text{error}|H_1] = P[X=0|H_1] + P[X=1|H_1] = 0.125 + 0.375 = 0.5 \)

Example: Constant Signal in Gaussian Noise

Hypotheses:
- \( H_0 \): Signal absent, \( X = w \) \( w \sim N(0,1) \)
- \( H_1 \): Signal present, \( X = v + w \) \( v \) = known constant (signal)

Likelihoods:
- \( f_{X|H_0} \sim N(0,1) \)
- \( f_{X|H_1} \sim N(v,1) \)

Log-Likelihood Ratio:

\[
\log(L(x)) = \log \left( \frac{f_{X|H_1}(x|H_1)}{f_{X|H_0}(x|H_0)} \right) = \frac{(x-v)^2}{2} + \frac{v^2}{2}
\]

ML Decision Rule: Select \( H_1 \) if \( x < v \)
Select \( H_0 \) if \( x > v \)

Error Probabilities:
- \( P[\text{error}|H_0] = P[X<v|H_0] = Q(v/2) \)
- \( P[\text{error}|H_1] = P[X>v|H_1] = Q(v/2) \)

Example: Radar

Hypotheses:
- \( H_0 \): Target is absent
- \( H_1 \): Target is present

Observation: \( X \) = radar echo. \( X \) is a Binomial RV \( (n, p) \)

Likelihoods:
- \( P[X|H_0] = \text{Binom}(n, 0) \)
- \( P[X|H_1] = \text{Binom}(n, p) \)

Likelihood matrix:

\[
\begin{array}{cccc}
X = 0 & X = 1 & X = 2 & X = 3 \\
P[X|H_0] & 0.00032 & 0.0064 & 0.0512 & 0.2048 & 0.4096 & 0.32768 \\
P[X|H_1] & 0.00032 & 0.0064 & 0.0512 & 0.2048 & 0.4096 & 0.32768 \\
\end{array}
\]

ML Decision Rule: Find and shade the largest entry in each column!  

Error Probabilities:
- \( P[\text{error}|H_0] = P[X=3|H_0] = 0.022 \)
- \( P[\text{error}|H_1] = P[X=2|H_1] = 0.00672 \)

Example: Localization in Sensor Networks

Hypotheses:
- \( H_0 \): Signal absent
- \( H_1 \): Signal present

Likelihoods:
- \( f_{X|H_0} \sim N(0,1) \)
- \( f_{X|H_1} \sim N(v,1) \)

Log-Likelihood Ratio:

\[
\log(L(x)) = \frac{(x-v)^2}{2} + \frac{v^2}{2}
\]

ML Decision Rule: Select \( H_1 \) if \( x < v \)
Select \( H_0 \) if \( x > v \)

Error Probabilities:
- \( P[\text{error}|H_0] = P[X<v|H_0] = Q(v/2) \)
- \( P[\text{error}|H_1] = P[X>v|H_1] = Q(v/2) \)

Example: Localization in Sensor Networks

Hypotheses:
- \( H_0 \): Signal absent
- \( H_1 \): Signal present

Likelihoods:
- \( f_{X|H_0} \sim N(0,1) \)
- \( f_{X|H_1} \sim N(v,1) \)

Log-Likelihood Ratio:

\[
\log(L(x)) = \frac{(x-v)^2}{2} + \frac{v^2}{2}
\]

ML Decision Rule: Select \( H_1 \) if \( x < v \)
Select \( H_0 \) if \( x > v \)

Error Probabilities:
- \( P[\text{error}|H_0] = P[X<v|H_0] = Q(v/2) \)
- \( P[\text{error}|H_1] = P[X>v|H_1] = Q(v/2) \)
8.4 Bayesian Hypothesis Testing

**MAP Decision Rule (discrete observation)**

For each \( X = x \), select the hypothesis that has the maximum posterior probability (MAP):  

\[
P[H|X = x] = \max_i \frac{P[X = x|H_i]P(H_i)}{P(x)} \quad \text{Bayes rule}
\]

- If \( P[H_i|X = x] > P[H_j|X = x] \), select \( H_i \), else select \( H_j \) (either if a tie)

Since \( P_x(x) \) is independent of \( H \), this is equivalent to selecting the hypothesis with the largest \( P[H_i, X = x] \).

- If \( P[H_1, X = x] > P[H_0, X = x] \), select \( H_1 \); else select \( H_0 \) (either if a tie)

Unlike the ML decision rule, the MAP decision rule depends on the prior probabilities \( P[H_i] \).

**Average probability of making a correct decision** is 

\[
P[\text{error}] = \sum_i P[\text{error}|H_i]P[H_i]
\]

**Example: Two Coins**

Let \( P[H_0] = 0.2, P[H_1] = 0.8 \)

**Likelihood matrix & ML decision rule**

<table>
<thead>
<tr>
<th></th>
<th>( X = 0 )</th>
<th>( X = 1 )</th>
<th>( X = 2 )</th>
<th>( X = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 )</td>
<td>0.612</td>
<td>0.384</td>
<td>0.006</td>
<td>0.008</td>
</tr>
<tr>
<td>( H_1 )</td>
<td>0.125</td>
<td>0.375</td>
<td>0.375</td>
<td>0.125</td>
</tr>
</tbody>
</table>

\[
P[\text{error}] = P[\text{error}|H_0]P[H_0] + P[\text{error}|H_1]P[H_1] = (0.125 + 0.375) \times 0.8 + (0.096 + 0.008) \times 0.2 = 0.4208
\]

**Joint Probability matrix & MAP decision rule**

<table>
<thead>
<tr>
<th></th>
<th>( X = 0 )</th>
<th>( X = 1 )</th>
<th>( X = 2 )</th>
<th>( X = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 )</td>
<td>0.1024</td>
<td>0.0768</td>
<td>0.0192</td>
<td>0.0016</td>
</tr>
<tr>
<td>( H_1 )</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

\[
P[\text{error}] = P[\text{error}|H_0]P[H_0] + P[\text{error}|H_1]P[H_1] = 0.1 \times (0.0768 + 0.0192 + 0.0016) = 0.1976
\]

**MAP decision rule** can be written in terms of the likelihood ratio \( L(x) \)

- If \( L(x) > T \), select \( H_1 \)
- If \( L(x) < T \), select \( H_0 \)

**Proof**

For any decision rule \( U(x) \),

\[
P[\text{error}] = P[U(x) = 0|H_0, X = x]P[H_0|X = x] + P[U(x) = 1|H_0, X = x]P[H_0|X = x]
\]

\[
P[\text{error}] = \sum_{x \in S_X} \left( P(U(x) = 0|H_0, X = x)P(H_0|X = x) + P(U(x) = 1|H_0, X = x)P(H_0|X = x) \right)
\]

For each \( x = X \), MAP decision rule selects smallest possible value for that term in the sum.