Hypothesis Testing

setup

Hypotheses: \( H_1 \),

Null hypothesis: \( H_0 \)

\[ P(H_1) = \Pi_1 , \quad P(H_0) = \Pi_0 . \]

Observations \( \rightarrow R \).

If you have no prior knowledge that one hypothesis is more likely than the other,

Form the Likelihood ratio \( \Lambda(R) \),

\[
\Lambda(R) = \frac{P(R | H_1)}{P(R | H_0)} \geq 1.
\]

If the LRatio \( \Lambda(R) \) is greater than one we declare \( R \) was generated under \( H_1 \), otherwise \( H_0 \).
$\Delta(r)$ \uparrow \quad \text{declare } H_i$

\[ \rightarrow \quad \text{declare } H_0 \]

\text{Example: } H_0: R \sim N(0, \sigma^2) \quad H_1: R \sim N(\mu, \sigma^2)

$\mathcal{P}(R|H_0)$ \quad \mathcal{P}(R|H_1)$

$R_0 \quad \sigma = ? \quad \mu$

- If $R$ lies in $R_0$, declare $H_0$ ($R < \sigma$)
- If $R$ lies in $R_1$, declare $H_1$ ($R > \sigma$)

What should $\sigma$ be?
What is \( \chi^2 \)?

For our example, \((H_i: R \sim N(\mu, \sigma^2)) \quad \text{and} \quad (H_0: R \sim N(0, \sigma^2))\):

\[
\Pr(R \mid H_i) = 1 \cdot \frac{e^{-\frac{1}{2\sigma^2} (x-\mu)^2}}{\sqrt{2\pi}\sigma^2}
\]

\[
\Pr(R \mid H_0) = \frac{1}{\sqrt{2\pi}\sigma^2} \cdot e^{-\frac{1}{2\sigma^2} (x^2)}
\]

\[
\Lambda(R) = \frac{e^{-\frac{1}{2\sigma^2} (x-\mu)^2}}{e^{-\frac{1}{2\sigma^2} x^2}} = e^{-\frac{1}{2\sigma^2} (x^2 - 2\mu x + \mu^2)}
\]

Set \( R = x \).

If our usual comparison is for \( \Lambda(R) \geq 1 \), we can simplify \( \Lambda(R) \) to get a sufficient statistic for \( R \),

\[
\Lambda(R) = e^{-\frac{1}{2\sigma^2} (\mu^2 - 2\mu x)} \geq 1
\]

\[-\frac{1}{2\sigma^2}(\mu^2 - 2\mu x) \geq \ln(1)\]
\[-\frac{1}{20^2} M^2 + \frac{MR}{5^2} \geq 0\]

\[\Rightarrow \frac{MR}{5^2} \geq \frac{M^2}{20^2}\]

\[\Rightarrow R \geq \frac{M}{2} = 8\]

\[H_0\]

\[H_1\]

Looking at our graph from before,

We say that

- if $R > \frac{M}{2}$
  \[\Rightarrow R \sim H_1\]
- if $R < \frac{M}{2}$
  \[\Rightarrow R \sim H_0\]

If $R = \frac{M}{2}$

- (Flip a coin)
This will minimize the total error,

\[ P(R < \theta | H_1) \Rightarrow \text{miss probability} \]

(say \( H_0 \) when it's actually \( H_1 \))

\[ P(R > \theta | H_0) \Rightarrow \text{false alarm probability} \]

(say \( H_1 \) when it's actually \( H_0 \))
What if the cost of a false alarm is higher than missing $H_i$? (e.g., saying $H_i$ when $H_0$ is true is worse than saying $H_0$ when $H_i$ is true)

\[
\begin{align*}
\text{(cost for false alarm) miss} \\
C_{10} : \text{Cost of deciding } H_i \text{ when } H_0 \text{ is true} \\
C_{10} : \text{Cost of deciding } H_0 \text{ when } H_i \text{ is true} \\
\text{(cost for misses)}
\end{align*}
\]

Instead

\[
P(R | H_i) C_{10} \overset{H_i}{\geq} P(R | H_0) C_{10} \overset{H_0}{\geq}
\]

Before $\Lambda(R) = \frac{P(R | H_i)}{P(R | H_0)} \overset{H_i}{\geq} 1 \overset{H_0}{\rightarrow}$ back to original case

Now, \[\Lambda(R) = \frac{P(R | H_i)}{P(R | H_0)} \overset{H_i}{\geq} \frac{C_{10}}{C_{01}} \text{ (if equal)}\]
Going back to the Gaussian example from before, (with costs)

\[
\frac{1}{2 \sigma^2} (2MR - \mu^2) \overset{H_1}{\geq} \ln \left( \frac{C_{10}}{C_{01}} \right) \overset{H_0}{\geq}
\]

\[
= \frac{1}{2 \sigma^2} \frac{2R}{\ln \left( \frac{C_{10}}{C_{01}} \right)} + \mu^2 \overset{H_0}{\geq}
\]

\[
\Rightarrow R \overset{H_1}{\geq} \frac{\sigma^2}{\mu} \ln \left( \frac{C_{10}}{C_{01}} \right) + \frac{\mu}{2} \overset{H_0}{=} \gamma
\]

\[
\downarrow
\]

This was zero before, so \( \gamma = \frac{\mu}{2} \).

If false alarms are 2x worse than misses, \( C_{10} = 2C_{01} \).
What about the converse?

\[ C_{o1} = 2C_{c0} \]

(misses are 2x worse than false alarms)

\[ R_{H_0} \geq \frac{M}{2} + \frac{\sigma^2}{M} (\ln\left(\frac{1}{2}\right)) = \gamma \]

\[ \uparrow \]

negative (shift to left)

\[ \gamma = \gamma_0 + \frac{\sigma^2}{M} (\ln\left(\frac{1}{2}\right)) \]
\[ R \geq \frac{M}{2} + \frac{\sigma^2}{M} \ln(2) = \delta \]

Positive (shifts \( \delta \) to right)

\[ \delta_{\text{new}} = \delta_0 + \frac{\sigma^2}{M} \ln(2) \]

Will more often declare \( \text{Ho} \) when \( \text{Hi} \) is true (more misses)

but, less area on other side,

hence we will have \underline{less} false alarms = just as we wanted!
Finally,

if we have prior knowledge than one hypothesis is more likely than another, we want to take this into account.

Maximum-A-Posteriori (MAP) RULE

\[ P(H_0) = \pi_0, \quad P(H_1) = \pi_1 \]

Looking at posterior prob (prob. that \( H_1 \) is true given the data)

\[ P(H_1|R) \geq P(H_0|R) \]

\[ H_0 \]
Choose the hypothesis that is most likely given the data,

\[ H_1 \]

\[ P(H_1 | R) \geq P(H_0 | R) \]

\[ H_0 \]

Using Bayes Rule,

\[ \frac{P(R | H_1) \cdot P(H_1)}{P(R)} \geq \frac{P(R | H_0) \cdot P(H_0)}{P(R)} \]

\[ H_0 \]

\[ \frac{P(R | H_1)}{P(R | H_0)} \geq \frac{P(H_0)}{P(H_1)} \]

MAP

if \( \frac{P(H_0)}{P(H_1)} = \frac{1}{2} \), the MAP rule is the same as likelihood ratio.