Nov 15

Hypothesis testing is like classification of machine learning.

Bayesian Decision Theory

some θ

$$\tilde{\theta} = \begin{cases} \theta^*, & \text{if } |\hat{\theta}_{MLE} - \theta^*| < 1/n \\ \hat{\theta}_{MLE}, & \text{if } |\hat{\theta}_{MLE} - \theta^*| > 1/n \end{cases}$$

Jake any estimator

Statistical Decision Theory

Example: Classifying Bernoulli's.

- $P_0 = 1$, $P_1 = 0$, $x = 1$

  Decision: p_0

- $P_0 = 0.99$, $p_1 = 0.1$, $x = \{1, 1, 1\}$

  Decision: p_0

$\mathcal{P} \xrightarrow{} $ family of distⁿ composing our model

$\mathcal{P} = \{P_\theta, \theta \in H\}$ — is the space of distributions.

* H_0 is the subset of "acceptable" elements of $\mathcal{P}$. (the hypothesis / null hypothesis)

* H_1 is the subset of "rejectable" elements of $\mathcal{P}$. (alternative / alternative hypothesis).

$\mathcal{P} = H_0 \cup H_1$; $H_0 \cap H_1 = \phi$

→ $\Theta = \Theta_0 \cup \Theta_1$ (parameter space).

$H_0 = \{P_\theta, \theta \subset \Theta_0\}$; $H_1 = \{P_\theta, \theta \subset \Theta_1\}$

* Data (one single obs): x, realization of r.v. X w. $dist^n$ P_θ, $\theta \in \Theta$

We want to construct $\delta(x) \in \{0, 1\}$, where $\delta(x)$ is the decision that $P_\theta \in H_0$ ($\delta(x)=0$) or not ($\delta(x)=1$) → Binary Form.

(The general form of this would be a randomized distribution of decision rule → This was Nash's decision.)

Objective Function (In ML: pattern recognition)

loss function of decision rule $L_\delta(\theta)$ → depends on θ that generates data.

→ Pattern recognition loss.

* Ex-ante problem: Before data is classified, we want to set the decision rule
* Dist^n is fully accessible.

Misclassification Probability

Misclassification Probability:

$$L_\delta(\theta) = \begin{cases} P_\theta(\delta(X) = 1), \text{ if } \theta \in \Theta_0 & : \text{Type 1 error} \\ P_\theta(\delta(X) = 0), \text{ if } \theta \in \Theta_1 & : \text{Type 2 error} \end{cases}$$

Going back ~~to~~ the Bernoulli Example:

$H_0 = \{\text{Bernoulli}(p_0)\}$ $H_1 = \{\text{Bernoulli}(p_1)\}$

$x \in \{0,1\}$

Options for decision rules :-

$$\delta^1(x) = \begin{cases} 1, & \text{if } x=1 \\ 0, & \text{otherwise if } x=0 \end{cases}$$

$$\delta^2(x) = \begin{cases} 0, & \text{if } x=1 \\ 1, & \text{if } x=0 \end{cases}$$

$$\delta^3(x) = \begin{cases} 1, & \text{if } x=1 \\ 1, & \text{if } x=0 \end{cases}$$

$$\delta^4(x) = \begin{cases} 0, & \text{if } x=1 \\ 0, & \text{if } x=0 \end{cases}$$

Reference

Probability

$$L_\delta(\theta) = \begin{cases} P_\theta(\delta(X)=1), \text{ if } \theta \in H_0 & : \text{Type 1 error} \\ P_\theta(\delta(X)=0), \text{ if } \theta \in H_1 & : \text{Type 2 error} \end{cases}$$

Decision Rule #	$L_\delta(\theta)$, $P_\theta \in H_0$	$L_\delta(\theta)$, $P_\theta \in H_1$
1	p_0	$1-p_1$
2	$1-p_0$	p_1
3	1	0
4	0	1

There is no global optimum decision rule in this case. (we want to min. loss