Lecture 12 - 9/19

Monday, September 19, 2022 7:04 AM

. Review

① Maximum likelihood estimator:

(i) distribution of random variable X belongs to a family of parametrized distributions

(ii) We are given an observation A, i.e., $X \in A$

(iii) We want to find the parameter that maximizes probability of A

② Inequalities:

Markov's inequality: $P(X \geq c) \leq \frac{E[X]}{c}$, $c > 0$ and X is a nonnegative random variable

Chebychev's inequality: $P(|X - E[X]| \geq a) \leq \frac{Var(X)}{a^2}$

③ Confidence interval:

. Suppose that each individual is in favor of policy A with probability p.

. We sample n individuals, X of them were in favor of policy A

. We do not know p.

(i) $\hat{p} = \frac{X}{n}$ is the point estimate of p

(ii) $P\left(p \in \left(\hat{p} - \frac{a}{2\sqrt{n}}, \hat{p} + \frac{a}{2\sqrt{n}}\right)\right) \geq 1 - \frac{1}{a^2}$

④ law of total probability and Bayes' formula:

. $P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i) P(A|E_i)$, E_is partitions Ω, $P(E_i) > 0$

. $P(E_i|A) = \dfrac{P(A|E_i)P(E_i)}{P(A|E_1)P(E_1) + \cdots + P(A|E_n)P(E_n)}$

⑤ law of total expectation:

. $E[g(X)] = \sum_i P(E_i) E[g(X)|E_i]$ E_is partitions Ω, $P(E_i) > 0$

. $E[g(X)|A] = \sum_i g(u_i) P(X = u_i | A)$

. Example: $Var(A|X) = E\left[(X - E(X|A))^2 | A\right] = E[X^2|A] - (E[X|A])^2$

. Today

• Today

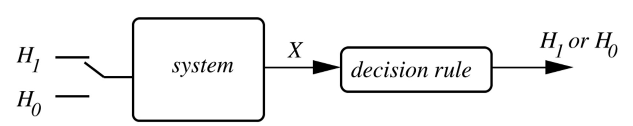
① Hypothesis testing

② ML detection rule and likelihood ratio test

---

① Hypothesis testing

- We have a system that generates random variables
- System is either in state H_0 or H_1
  - if in H_0, pmf of X is given by p_0, $p_0(k) = P(X=k \mid H_0)$
  - if in H_1, pmf of X is given by p_1, $p_1(k) = P(X=k \mid H_1)$
- We observe an outcome from the system, i.e., we observe $\{X=m\}$
- There is a decision rule that assigns hypothesis to each outcome



example: in above figure system is in state H_1.

① • System generates photos:

H_0: photo is real

H_1: photo is fake

- We observe an outcome, i.e., photo generated by system
- Based on decision rule, we decide photo is fake or not.

② • System generates brain scans

H_0: there is no tumor

H_1: there is a brain tumor

- We observe one scan of brain generated by the system
- Based on decision rule, we decide whether there is a tumor or not

Def. likelihood matrix: suppose that X is a random variable with values in $\{1,2,3,4\}$

	$X=1$	$X=2$	$X=3$	$X=4$	
H_0	$P(X=1 \mid H_0)$	$P(X=2 \mid H_0)$	$P(X=3 \mid H_0)$	$P(X=4 \mid H_0)$	⟶ pmf of X given H_0 is true

	X=1	X=2	X=3	X=4	
H_0	$P(X=1 \mid H_0)$	$P(X=2 \mid H_0)$	$P(X=3 \mid H_0)$	$P(X=4 \mid H_0)$	$\rightsquigarrow$ pmf of X given H_0 is true
H_1	$P(X=1 \mid H_1)$	$P(X=2 \mid H_1)$	$P(X=3 \mid H_1)$	$P(X=4 \mid H_1)$	$\rightsquigarrow$ pmf of X given H_1 is true

example:

	X=1	X=2	X=3	X=4
H=0	0.1	0.2	0.3	0.4
H=1	0.5	0	0.2	0.3

Def: decision rule: which hypothesis to pick after each observation

	X=1	X=2	X=3	X=4	
H=0	$P(X=1 \mid H_0)$	$\boxed{P(X=2 \mid H_0)}$	$P(X=3 \mid H_0)$	$P(X=4 \mid H_0)$	$\rightsquigarrow$ pick H_0 if $X=2$
H=1	$\boxed{P(X=1 \mid H_1)}$	$P(X=2 \mid H_1)$	$\boxed{P(X=3 \mid H_1)}$	$\boxed{P(X=4 \mid H_1)}$	$\rightsquigarrow$ pick H_1 if $X=1$ or 3 or 4

- If $X=2$, decision rule declares H_0
- If $X \in \{1,3,4\}$, decision rule declares H_1

example:

	X=1	X=2	X=3	X=4
H=0	0.1	$\boxed{0.2}$	$\boxed{0.3}$	0.4
H=1	$\boxed{0.5}$	0	0.2	$\boxed{0.3}$

- If $X \in \{2,3\}$, decision rule declares H_0
- If $X \in \{1,4\}$, decision rule declares H_1

Def: miss & false alarm

There are four possibility of our hypothesis:

(i) H_0 true, H_1 declared $\rightsquigarrow$ correct guess

(ii) H_1 true, H_1 declared $\rightsquigarrow$ correct guess

(iii) H_0 true, H_1 declared $\rightsquigarrow$ false alarm

(iv) H_1 true, H_0 declared $\rightsquigarrow$ miss

$$P_{\text{false alarm}} = P(H_1 \text{ declared} \mid H_0 \text{ is true})$$

$$P_{\text{miss}} = P(H_0 \text{ declared} \mid H_1 \text{ is true})$$

Give the following likelihood matrix

	$X=1$	$X=2$	$X=3$	$X=4$
$H=0$	$P(X=1\|H_0)$	$\boxed{P(X=2\|H_0)}$	$P(X=3\|H_0)$	$P(X=4\|H_0)$
$H=1$	$\boxed{P(X=1\|H_1)}$	$P(X=2\|H_1)$	$\boxed{P(X=3\|H_1)}$	$\boxed{P(X=4\|H_1)}$

We have,

$$P_{\text{false alarm}} = P(X=1|H_0) + P(X=3|H_0) + P(X=4|H_0)$$

$$P_{\text{miss}} = P(X=2|H_1)$$

example:

	$X=1$	$X=2$	$X=3$	$X=4$
$H=0$	0.1	$\boxed{0.2}$	$\boxed{0.3}$	0.4
$H=1$	$\boxed{0.5}$	0	0.2	$\boxed{0.3}$

$$P_{\text{false alarm}} = 0.1 + 0.4$$

$$P_{\text{miss}} = 0.2$$

Important remarks.

- $P_{\text{false alarm}}$ and P_{miss} are conditional probabilities
- $P(\text{wrong decision}) = P(\{\text{system in } H_0\} \cap \{\text{decision in } H_1\}) + P(\{\text{system in } H_1\} \cap \{\text{decision in } H_0\})$

$$= P(\{\text{system in } H_0\})\, P(\{\text{decision in } H_1\} \mid \{\text{system in } H_0\}) + P(\{\text{system in } H_1\})\, P(\{\text{decision in } H_0\} \mid \{\text{system in } H_1\})$$

$$= P(\{\text{system in } H_0\}) \cdot P_{\text{false alarm}} + P(\{\text{system in } H_1\}) \cdot P_{\text{miss}}$$

Which decision rule should we pick?

---

② Maximum likelihood detection rule and likelihood ratio test

Pick the most likely outcome, i.e., the one that maximizes the likelihood:

In each column of likelihood matrix, pick the largest value

Pick the most likely outcome, i.e., the one that maximizes the likelihood:

- In each column of likelihood matrix, pick the largest value
- If entries of column are identical, any choice is OK.

example:

	$X=1$	$X=2$	$X=3$	$X=4$
$H=0$	0.1	$\boxed{0.2}$	$\boxed{0.3}$	$\boxed{0.4}$
$H=1$	$\boxed{0.5}$	0	0.2	0.3

Def: likelihood ratio test: $\Lambda(k) = \dfrac{P_1(k)}{P_0(k)} = \dfrac{P(X=k \mid H_1)}{P(X=k \mid H_0)}$

$$\Lambda(X): \begin{cases} \Lambda(X) > 1, & \text{declare } H_1 \\ \Lambda(X) < 1, & \text{declare } H_0 \\ \Lambda(X) = 1, & \text{either } H_0 \text{ or } H_1 \end{cases}$$

Important remark:

- ML detection rule is same as maximum likelihood estimator:

  detection rule is H_i if $i = \arg\max_{j \in \{0,1\}} P(X=k \mid H_j) = \arg\max_{j\in\{0,1\}} P_j(k)$

2.35. **[Matching Poisson means]**
Consider hypotheses H_0 and H_1 about a two dimensional observation vector $X = (X_1, X_2)$. Under H_0, X_1 and X_2 are mutually independent, and both have the Poisson distribution with mean 4. Under H_1, X_1 and X_2 are mutually independent, X_1 has the Poisson distribution with mean 2, and X_2 has the Poisson distribution with mean 6.

(a) Describe the ML rule for H_0 vs. H_1. Display your answer by indicating how to partition the set of possible observations, $\{(i, j) : i \geq 0, j \geq 0\}$, into two sets, Γ_0 and Γ_1, for which the decision is H_0 if $(X_1, X_2) \in \Gamma_0$ and H_1 if $(X_1, X_2) \in \Gamma_1$.

H_0: X_1 & X_2 are indep, and both are Poi(4)

H_1: X_1 & X_2 are indep, and X_1 is Poi(2), X_2 is Poi(6)

$$P(X_1=m, X_2=n \mid H_0) = P(X_1=m \mid H_0)\, P(X_2=n \mid H_0) = e^{-4}\frac{4^m}{m!} \times e^{-4}\frac{4^n}{n!}$$

$$P(X_1=m, X_2=n \mid H_0) = P(X_1=m \mid H_0)\,P(X_2=n \mid H_0) = e \,\frac{4}{m!} \times e \,\frac{4}{n!}$$

$$P(X_1=m, X_2=n \mid H_1) = P(X_1=m \mid H_1)\,P(X_2=n \mid H_1) = e^{-2}\frac{2^m}{m!} \times e^{-6}\frac{6^n}{n!}$$

if we observe $X_1 = k$ and $X_2 = \ell$:

$$\Lambda(X_1=k, X_2=\ell) = \frac{e^{-2}\frac{2^k}{k!} \times e^{-6}\frac{6^\ell}{\ell!}}{e^{-4}\frac{4^k}{k!} \times e^{-4}\frac{4^\ell}{\ell!}} = \frac{2^k \times 6^\ell}{4^k \times 4^\ell} = \frac{3^\ell}{2^k \times 2^\ell}$$

We declare H_0 if $\Lambda(k,\ell) > 1$

$$\Leftrightarrow 3^\ell > 2^{\ell+k}$$

$$\Leftrightarrow \ell \ln 3 > (\ell+k)\ln 2$$

$$\Leftrightarrow \ell \ln \tfrac{3}{2} > k \ln 2$$

$$\Leftrightarrow \ell > k \times \frac{\ln 2}{\ln 1.5}$$