Lecture 13 - 9/21

Tuesday, September 20, 2022 8:21 AM

Review

① Hypothesis testing

(i) The system generates random variable X

- If the system is in state H_0, the pmf of X is given by p_0
- If the system is in state H_1, the pmf of X is given by p_1

$$p_0(k) = P(X=k \mid H_0) \quad \text{and} \quad p_1(k) = P(X=k \mid H_1)$$

(ii) We observe a realization of X, i.e., we observe $X=m$.

(iii) We guess the state of system, using our observation, based on the decision rule

Def. Likelihood matrix: each row corresponds to a hypothesis

each column corresponds to an observation

each circled value (or underlined values) is a decision rule for the column

Def. probability of miss, p_{miss}: probability of guessing H_0 given system is in state H_1,

$$p_{miss} = P(\text{Decision rule returns } H_0 \mid \text{system is in } H_1)$$

Def. probability of false alarm, $p_{false\ alarm}$: probability of guessing H_1 given system is in state H_0,

$$p_{false\ alarm} = P(\text{Decision rule returns } H_1 \mid \text{system is in } H_0)$$

Today:

① Maximum likelihood (ML) decision rule

② Threshold decision rule

③ Maximum a posteriory probability (MAP) decision rule

---

① ML decision rule:

From viewpoint of Maximum likelihood estimation, we have two pmfs and we want to pick the one that maximizes likelihood. Recall that $p_0(k)$ is the likelihood of observing $X=k$ given the system is in state H_0. ML decision rule maximizes the likelihood of our observation.

- Given $X=k$, $\begin{cases} \text{declare } H_0 \text{ if } p_0(k) > p_1(k) \\ \text{declare } H_1 \text{ if } p_0(k) < p_1(k) \\ \text{declare either if } p_0(k) = p_1(k) \end{cases}$

- In likelihood matrix, pick the largest element in each column.

Def: likelihood ratio test: define $\Lambda(k) = \dfrac{p_1(k)}{p_0(k)} = \dfrac{P(X=k|H_1)}{P(X=k|H_0)}$

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

② Threshold decision rule

We can generalize the likelihood ratio test by introducing a threshold τ:

$$\Lambda(k): \begin{cases} \text{declare } H_0 \text{ if } & \Lambda(k) < \tau \\ \text{declare } H_1 \text{ if } & \Lambda(k) > \tau \\ \text{declare either if } & \Lambda(k) = \tau \end{cases}$$

In particular, ML decision rule is a threshold rule with $\tau = 1$.

③ Maximum a posteriori probability

Suppose that we have a prior π_0 and π_1.

$\pi_0 = P(\text{System in } H_0)$, $\pi_1 = P(\text{System in } H_1)$ $\rightsquigarrow$ $\boxed{\text{prior probabilities}}$

These are our guess about the state of the system before (prior to) any observation.

After observing a realization $X=k$, the probability that system is in state H_0 is given by:

$$P(\text{system in } H_0 | X=k) = \frac{P(H_0, X=k)}{P(X=k)} = \frac{P(H_0)\cdot P(X=k|H_0)}{P(H_0)\cdot P(X=k|H_0) + P(H_1)\cdot P(X=k|H_1)} = \frac{\pi_0 \cdot p_0(k)}{\pi_0\cdot p_0(k) + \pi_1 \cdot p_1(k)}$$

Similarly

$$P(\text{system in } H_1 | X=k) = \frac{\pi_1 \cdot p_1(k)}{\pi_0 \cdot p_0(k) + \pi_1 \cdot p_1(k)}$$

$\boxed{\text{a posteriori probabilities}}$

Convention: $P(A,B) = P(AB) = P(A\cap B)$

The MAP decision rule make guess by maximizing probability of state given the observation.

- Given $X=k$ $\begin{cases} \text{declares } H_0 \text{ if } P(H_0|X=k) > P(H_1|X=k) \\ \text{declares } H_1 \text{ if } P(H_0|X=k) < P(H_1|X=k) \\ \text{declares eithe if } P(H_0|X=k) = P(H_1|X=k) \end{cases}$ OR equivalently $\begin{cases} \text{declares } H_0 \text{ if } P(H_0, X=k) > P(H_1, X=k) \\ \text{declares } H_1 \text{ if } P(H_0, X=k) < P(H_1, X=k) \\ \text{declares eithe if } P(H_0, X=k) = P(H_1, X=k) \end{cases}$

• Notice that the decision rule depends on the ratio $\frac{P(H_1|X=k)}{P(H_0|X=k)}$

$$\frac{P(H_1|X=k)}{P(H_0|X=k)} = \frac{\pi_1 \cdot p_1(k)}{\pi_0 \cdot p_0(k)} = \frac{\pi_1}{\pi_0} \times \Lambda(k)$$

Hence, MAP decision rule is a threshold rule with $\tau = \frac{\pi_0}{\pi_1}$

$$\Lambda(k): \begin{cases} \text{declare } H_0 \text{ if } & \Lambda(k) < \frac{\pi_0}{\pi_1} \\ \text{declare } H_1 \text{ if } & \Lambda(k) > \frac{\pi_0}{\pi_1} \\ \text{declare either if } & \Lambda(k) = \frac{\pi_0}{\pi_1} \end{cases}$$

Important Remark:

Instead of looking at $\Lambda(k) = \frac{p_1(k)}{p_0(k)}$, we can check the value of $\log \Lambda(k) = \log p_1(k) - \log p_0(k)$

• For ML decision rule: $\log \Lambda(k) = \log p_1(k) - \log p_0(k): \begin{cases} \text{declare } H_0 \text{ if } & \log \Lambda(k) < 0 \\ \text{declare } H_1 \text{ if } & \log \Lambda(k) > 0 \\ \text{declare either if } & \log \Lambda(k) = 0 \end{cases}$

• For MAP decision rule: $\log \Lambda(k) = \log p_1(k) - \log p_0(k): \begin{cases} \text{declare } H_0 \text{ if } & \log \Lambda(k) < \log(\frac{\pi_0}{\pi_1}) \\ \text{declare } H_1 \text{ if } & \log \Lambda(k) > \log(\frac{\pi_0}{\pi_1}) \\ \text{declare either if } & \log \Lambda(k) = \log(\frac{\pi_0}{\pi_1}) \end{cases}$

• As long as you use the same logarithm base in above (on all side of inequalities) the result is same.

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)

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

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

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

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

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

$$\Rightarrow \log \Lambda(X_1 = k, X_2 = \ell) = \ell \times \log 3 - (k+\ell) \times \log 2 = \ell \times \log 1.5 - k \times \log 2$$

$$\text{ML}: \begin{cases} \text{declare } H_0 \text{ if } & \ell \times \log 1.5 - k \times \log 2 < 0 \\ \text{declare } H_1 \text{ if } & \ell \times \log 1.5 - k \times \log 2 > 0 \\ \text{declare either if } & \ell \times \log 1.5 - k \times \log 2 = 0 \end{cases}$$

(b) Describe the MAP rule for H_0 vs. H_1, assuming the prior distribution with $\frac{\pi_0}{\pi_1} = 2$. 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 $(X_1, X_2) \in \Gamma_1$.

Notice that $\frac{\pi_0}{\pi_1} = 2$. Using logarithim base 2 in the above formula, we have

$$\text{MAP}: \begin{cases} \text{declare } H_0 \text{ if } & \ell \times \log 1.5 - k < 1 \\ \text{declare } H_1 \text{ if } & \ell \times \log 1.5 - k > 1 \\ \text{declare either if } & \ell \times \log 1.5 - k = 1 \end{cases}$$