Biostat 802 Homework 3

David (Daiwei) Zhang

February 28, 2018

1 Discrete Distributions

We first find the distribution of the likelihood ratios for P_1 and P_2 against P_0 under the measure of P_0. We have

L_{P_1}	0.85	2.00		2.50	4.00	∞
$P_0(L_{P_1})$	0.92	0.04		0.02	0.02	0.00
x	6	1	4	2	3	5
$P_0(x)$	0.92	0.03	0.01	0.02	0.02	0.00

L_{P_2}	0.00	0.78	2.50	3.00	6.00	∞
$P_0(L_{P_2})$	0.01	0.92	0.02	0.03	0.02	0.00
x	4	6	2	1	3	5
$P_0(x)$	0.01	0.92	0.02	0.03	0.02	0.00

where $L_{p_1} = P_1(X)/P_0(X)$ and $L_{p_2} = P_2(X)/P_0(X)$.

Now in order to find a level-α test, we pick the observations with the highest likelihood ratio (the right most side of the two tables above).

1. $\alpha = 0.01$. For P_1, we should reject $X = 5$ and $\frac{1}{2}$ of $X = 3$, in order to achieve maximum power. By looking at the table for P_2, we find that power under P_2 is maximized by the same rejection rule. Thus the UMP test exists, which is

$$\phi(x) = \begin{cases} 0, & \text{if } x \in \{1, 2, 4, 6\}, \\ \frac{1}{2}, & \text{if } x = 3, \\ 1, & \text{if } x = 5 \end{cases}$$

2. $\alpha = 0.05$. The situation is different here. For P_1, we must reject $X \in \{2, 3, 5\}$ and $\frac{1}{4}$ of $X \in \{1, 4\}$ to achieve maximum power. But for P_2, we must reject $X \in \{1, 3, 5\}$ and nothing else, since $P_0[X \in \{1, 3, 5\}] = 0.05$ already. Thus a UMP test does not exist.

3. $\alpha = 0.07$. Now as we increase α, the situation changes again. This time for P_2, we must reject $X \in \{2, 1, 3, 5\}$ exactly to achieve maximum power. Luckily, this is allowed by the UMP rejection rule for P_1, too. For P_1, we must reject $X \in \{2, 3, 5\}$ with probability 1 and $X \in \{1, 4\}$ with probability $\frac{3}{4}$. Since $P_0[X = 1] = 0.03$ and $P_0[X = 4] = 0.01$, things work out exactly well if we reject $X = 1$ but not $X = 3$. Overall, the UMP test is

$$\phi(x) = \begin{cases} 0, & \text{if } x \in \{4, 6\}, \\ 1, & \text{if } x \in \{1, 2, 3, 5\} \end{cases}$$

2 Bayes Hypothesis Testing

2.1 Probability of Error

Let $\delta(x)$ be our decision rule. Then the overall probability of error is

$$\begin{aligned} P[I(H_i)(1 - \delta(X)) = 1] =& E[I(H_i)(1 - \delta(X))] \\ =& \pi_0 E\Big[\delta(X)\Big| I(H_0)\Big] + \pi_1 E\Big[1 - \delta(X)\Big| I(H_1)\Big] \\ =& \pi_0 E\Big[E[\delta(X)]\Big| I(H_0)\Big] + \pi_1 E\Big[E[1 - \delta(X)]\Big| I(H_1)\Big] \\ =& \pi_0 E\Big[\psi(X)\Big| I(H_0)\Big] + \pi_1 E\Big[1 - \psi(X)\Big| I(H_1)\Big] \\ =& \pi_0 E_0 \psi(X) + \pi_1 E_1[1 - \psi(X)] \end{aligned}$$

2.2 Bayes Test

We have

$$\begin{aligned} P[\text{Error}] =& \pi_0 E_0 \psi(X) - \pi_1 E_1[\psi(X)] + \pi_1 \\ =& \pi_0 \Big\{ E_0 \psi(X) - \frac{\pi_1}{\pi_0} E_1[\psi(X)] \Big\} + \pi_1 \end{aligned}$$

In order to minimize $P[\text{Error}]$, we just need to find the $\psi(x)$ that minimizes

$$\begin{aligned} E_0\psi(X) - \frac{\pi_1}{\pi_0}E_1[\psi(X)] &= \int_{x\in\mathcal{X}} \left[P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)\right]\psi(x)dx \\ &= \int_{S_-} \left[P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)\right]\psi(x)dx \\ &+ \int_{S_+} \left[P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)\right]\psi(x)dx. \end{aligned}$$

Here dx is the Lebesgue measure, and S_- and S_+ corresponds to the subset of $\mathcal{X}$ where $P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)$ are negative and positive, respectively. (We don't need to worry about the case when this value is zero, since in that case the integral would be zero, too.) Since S_- and S_+ are disjoint, we just need to minimize the integrals over them separtely. By viewing ψ as a weight function ranging between 0 and 1, we see that

$$\int_{S_-} \left[P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)\right]\psi(x)dx$$

is minimized by setting $\psi(x)$ to 1 over all of S_-, and that

$$\int_{S_+} \left[P_0(x) - \frac{\pi_1}{\pi_0}P_1(x)\right]\psi(x)dx$$

is minimized by setting $\psi(x)$ to 0 over all of S_+. But S_- is exactly the region where

$$\frac{P_1(x)}{P_0(x)} > \frac{\pi_0}{\pi_1}.$$

Thus the critical function for minimizing the probability of error is

$$\psi(x) = \begin{cases} 0, & \text{if } \frac{P_1(x)}{P_0(x)} < \frac{\pi_0}{\pi_1} \\ \text{anything}, & \text{if } \frac{P_1(x)}{P_0(x)} = \frac{\pi_0}{\pi_1} \\ 1, & \text{if } \frac{P_1(x)}{P_0(x)} > \frac{\pi_0}{\pi_1} \end{cases}$$

2.3 Posterior Probability and Likelihood Ratio

Consider the case when H_1 has a larger posterior probability than H_0, that is,

$$\frac{\pi_1 P_1(x)}{\pi_0 P_0(x) + \pi_1 P_1(x)} > \frac{\pi_0 P_0(x)}{\pi_0 P_0(x) + \pi_1 P_1(x)},$$

which is exactly

$$\frac{P_1(x)}{P_0(x)} > \frac{\pi_0}{\pi_1},$$

the creteria for the likelihood ratio test. The same argument applies to the case when H_0 has a larger posterior probability than H_1 and the case when H_0 and H_1 have the same posterior probability.

3 Power of Sufficient Statistic

For any θ, the power $\psi(T(X))$ is

$$\begin{aligned} E_\theta[\psi(T(X))] =& E_\theta\Big[E[\psi(X)|T(X)]\Big] \\ =& E_\theta[\psi(X)], \end{aligned}$$

which is the power of $\psi(X)$.