CSCI 4100 Fall 2018
Assignment 3 Answers

Damin Xu
661679187

September 24, 2018

Exercise 1.13

(a) $P_1 = \mu\lambda$
$P_2 = (1-\lambda)(1-\mu)$
$P = P_1 + P_2 = \mu\lambda + (1-\lambda)(1-\mu) = 1 + 2\mu\lambda - \mu - \lambda$

(b) Since $P = 1 + 2\mu\lambda - \mu - \lambda$:
$$P = 1 - \lambda + (2\lambda - 1)\mu$$
We want h be independent of μ, so $2\lambda - 1 = 0$.
Therefore, $\lambda = 1/2$

Exercise 2.1

(1) For positive rays, break point is when k = 2.
Use the formula $m_H(N) = N + 1$, $m_H(2) = 3 < 2^2 = 4$ So it is true $m_H(k) < 2^k$ at break point.

(2) For positive intervals, $m_H(N) = \frac{1}{2}N^2 + \frac{1}{2}N + 1$.
Compute:
$$\frac{1}{2}N^2 + \frac{1}{2}N + 1 < 2^N$$
Get:$k = 3$
$m_H(3) = 7 < 8 = 2^3$ So it is true $m_H(k) < 2^k$ at break point.

(3) Because for convex sets, $m_H(N) = 2^N$, so there is no break point.

Exercise 2.2

(a) **Theorem 2.4.** If $m_H(k) < 2^k forsomevaluek, then:$

$$m_H(N) \leq \sum_{i=0}^{k-1} \binom{N}{i}$$

for all N. The RHS is polynomial in N of degree $k-1$

(i) For positive rays, break point is k = 2, and $m_H(2) = 3$, so

$$\begin{aligned} m_H(N) &= N+1 \\ &\leq \sum_{i=0}^{k-1} \binom{N}{i} \\ &= \sum_{i=0}^{1} \binom{N}{i} \\ &= 1+N \end{aligned}$$

(ii) For positive intervals, break point is k = 3, so

$$\begin{aligned} m_H(N) &= \frac{1}{2}N^2 + \frac{1}{2}N + 1 \\ &\leq \sum_{i=0}^{k-1} \binom{N}{i} \\ &= \sum_{i=0}^{2} \binom{N}{i} \\ &= 1 + N + \frac{1}{2}N(N-1) \\ &= \frac{1}{2}N^2 + \frac{1}{2}N + 1 \end{aligned}$$

(iii) For convex sets, there is no break point, but we can still assume $k = N+1$, so

$$\begin{aligned} m_H(N) &= 2^N \\ &\leq \sum_{i=0}^{k-1} \binom{N}{i} \\ &= \sum_{i=0}^{N+1-1} \binom{N}{i} \\ &= 2^N \end{aligned}$$

(b) There cannot be a hyothesis set which $m_H(N) = N + 2^{[N/2]}$, because $N + 2^{[N/2]}$ is an exponential function and $\sum_{i=0}^{k-1} \binom{N}{i}$ is kind of polynomials function, so there is no k achieving

$$N + 2^{[N/2]} \leq \sum_{i=0}^{k-1} \binom{N}{i}$$

Exercise 2.3
According to the formula

$$d_{vc} = k - 1$$

(i) $d_{vc} = k - 1 = 2 - 1 = 1$

(ii) $d_{vc} = k - 1 = 3 - 1 = 2$

(iii) $d_{vc} = k - 1 = \infty - 1 = \infty$

Exercise 2.6

(a) Beacuse $E_{out}(g) \leq E_{in}(g) + \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$, and $M = 1000 N_{in\_g} = 400 N_{test\_g} = 200 \delta = 0.05$.
$E_{in}(g) = \sqrt{\frac{1}{800} ln \frac{2000}{0.05}}$
$E_{test}(g) = \sqrt{\frac{1}{400} ln \frac{2000}{0.05}}$
So $E_{test}(g) \geq E_{in}(g)$.

$E_{test}(g)$ has a higher error bar.

(b) If there are too many examples for testing, the test set will be smaller and the result could be worse.

Problem 1.11
Suppose there are N examples, then for supermarkets,

$$E_{in} = \frac{1}{N} \sum_{i=1}^{n} (10 \times [h(x_n) \neq 1] + [h(x_n) \neq -1])$$

For CIA,

$$E_{in} = \frac{1}{N} ([h(x_n) \neq 1] + 1000 \times [h(x_n) \neq -1])$$

Problem 1.12

(a) Because we need to find the stationary point, we have to compute $E_in''(h)$ first.

$$E_{in}(h) = \sum_{n=1}^{N}(h-y_n)^2$$
$$E'_{in}(h) = \sum_{n=1}^{N}[2(h-y_n)]$$
$$E''_{in}(h) = \sum_{n=1}^{N}(2) = 2N$$

so $E'_{in}(h_{mean}) = 0$,

$$h_{mean} = \frac{1}{N}\sum_{n=1}^{N} y_n$$

(b)

$$F(a) = \int_{-\infty}^{a}(a-x)f(x)dx + \int_{a}\infty(x-a)f(x)dx$$
$$F'(a) = \int_{-\infty}^{a} f(x)dx - \int_{a}\infty f(x)dx$$
$$F''(a) = 2f(a)$$

Because $F''(a) > 0$, we want minimun from $F'(a) = 0$.
Then $a = x_{mid}$.
Let $P(y = y_i) = \frac{1}{N}(i = 1, 2, 3...N)$, then $F(h) = \frac{1}{N}E_{in}(h) = \frac{1}{N}\sum_{n=1}^{N}|h - y_n|$
So when $h = y_{med}, E_{in}(h)$ is the minimum.

(c) As y_N becomes as an outlier, h_{mean} becomes more and more close to ∞, but h_{med} does not change.