Final

Lei Xia (MML)

December 20, 2020

1. (a) Given that $f : (0,1] -> 0,1, f(x) = Z_i, \mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = 0) = 1/2$, it is known that $\mathbb{P}(f(x) = 0) = \mathbb{P}(f(x) = 1) = 1/2$. For f to be continuous, it cannot jump between 0 and 1: should always be 0 or 1. For $i \in [1, m]$, if $Z_1 = 0$, then $Z_2, ..., Z_m$ should all be 0 $=>$ $P1 = \frac{1}{2} * (\frac{1}{2})^{m-1}$; if $Z_1 = 1$, then $Z_2, ..., Z_m$ should all be 1 $=>$ $P2 = \frac{1}{2} * (\frac{1}{2})^{m-1}$. The probability that f is continuous is $P1 + P2 = (\frac{1}{2})^{m-1}$.

   (b) Since X_i are IID independently distributed on $(0, 1]$, $\mathbb{P}(X_1 \in (\frac{i-1}{m}, \frac{i}{m}]) = \frac{i}{m} - \frac{i-1}{m} = \frac{1}{m}$

   (c) $\mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{i-1}{m}, \frac{i}{m}] \neq \varnothing) = 1 - \mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{i-1}{m}, \frac{i}{m}] = \varnothing) = 1 - \mathbb{P}(X_1 \cap (\frac{i-1}{m}, \frac{i}{m}] = \varnothing) * \mathbb{P}(X_2 \cap (\frac{i-1}{m}, \frac{i}{m}] = \varnothing) ... * \mathbb{P}(X_n \cap (\frac{i-1}{m}, \frac{i}{m}] = \varnothing) = 1 - (1 - \frac{1}{m})^n$

   (d) $(\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}] = (\frac{f(X_{n+1})-1}{m}, \frac{f(X_{n+1}}{m}]$. Also, $f(x) = Z_i, \mathbb{P}(Z_i = 1) = \mathbb{P}(Z_i = 0) = 1/2$, so $f(X_{n+1} = 0) = f(X_{n+1} = 1) = 1/2$.

      i. $\mathbb{P} = 1/2 : f(X_{n+1}) = 0$, then $(\frac{f(X_{n+1})-1}{m}, \frac{f(X_{n+1})}{m}] = (-\frac{1}{m}, 0]$. Given $X_i \in (0, 1]$, $\mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}] = \varnothing) = \mathbb{P}(\{X_1, ..., X_n\} \cap (-\frac{1}{m}, 0] = \varnothing) = 1 => \mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}] \neq \varnothing) = 0$

      ii. $\mathbb{P} = 1/2 : f(X_{n+1}) = 1 => (\frac{f(X_{n+1})-1}{m}, \frac{f(X_{n+1})}{m}] = (0, \frac{1}{m}] => \mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}] \neq \varnothing) = \mathbb{P}(\{X_1, ..., X_n\} \cap (0, \frac{1}{m}] \neq \varnothing) = 1 - \mathbb{P}(\{X_1, ..., X_n\} \cap (0, \frac{1}{m}] = \varnothing) = (c) = 1 - (1 - \frac{1}{m})^n$.

      As a result, $\mathbb{P}(\{X_1, ..., X_n\} \cap (\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}] \neq \varnothing) = \frac{1}{2}[1 - (1 - \frac{1}{m})^n]$.

   (e) Y_{n+1} needs to be Y_i all the time, which means that $X_i \in (\frac{[mX_{n+1}]-1}{m}, \frac{[mX_{n+1}]}{m}]$ always holds for X_i, i ranging from $[1, n]$. $\mathbb{P}(\hat{Y}_{n+1} = Y_i) = \sum_{k=1}^{n} \frac{1}{2}[1 - (1 - \frac{1}{m})^n] = \frac{1}{2}n - (m-1)[1 - (1 - \frac{1}{m})^n]$.

2. (a) For period 1, when $\omega R_1 \geq 1, C_1 = 1 + \alpha(\omega R_1 - 1)$; when $\omega R_1 \leq 1, C_1 = 1 + \beta(\omega R_1 - 1)$. Combine these two factors into one, I get $C_1 = \omega R_1 = 1 + \alpha(\omega R_1 - 1)^+ - \beta(\omega R_1 - 1)^-$. Given that the return in each period is independent from other periods', then

   $$C_n = (\omega R_1)^n => ln(C_n) = n \cdot ln(\omega R_1) => \lim_n \frac{1}{n} ln(C_n) = ln(\omega R_1) = ln[1 + \alpha(\omega R_1 - 1)^+ - \beta(\omega R_1 - 1)^-].$$

   (b) With probability $p1$,

   $$f_1(\omega) = \omega(1 + r_1) + (1 - \omega)[1 + \alpha(\omega R_1 - 1)^+ - \beta(\omega R_1 - 1)^-];$$

   and with probability $(1 - p_1)$,

   $$f_2(\omega) = \omega(1 + r_2) + (1 - \omega)[1 + \alpha(\omega R_2 - 1)^+ - \beta(\omega R_2 - 1)^-].$$

   Given $(R_2 - 1) \geq 0$ and $(R_2 - 1) \leq 0$, then

   $$f_1(\omega) = \omega(1 + r_1) + (1 - \omega)[1 + \alpha(\omega r_1 - 1), \quad f_2(\omega) = \omega(1 + r_2) + (1 - \omega)[1 + \beta(\omega r_2 - 1)],$$

   then

   $$f(\omega) = p_1 \cdot f_1(\omega) + p_2 \cdot f_2(\omega)$$
   $$= -(\alpha p_1 r_1 + \beta(1 - p_1) r_2)\omega^2 + [(\alpha(r_1 + 1) - \beta(r_2 + 1) + r_1 - r_2)p_1 + \beta(r_2 + 1) + r_2]\omega - (\alpha - \beta)p_1 + (1 - \beta).$$

Standard quadratic equation. To max $f(\omega)$,

$$\omega = \frac{[\alpha(r_1+1) - \beta(r_2+1) + (r_1 - r_2)]p_1 + \beta(r_2+1) + r_2}{2(\alpha p_1 r_1 + \beta(1-p_1)r_2)}$$

When $r_1 = 1, r_2 = -0.5, p_1 = 0.5$,

$$\omega = \frac{4\alpha + \beta + 1}{4\alpha - 2\beta}.$$

i.

$$\alpha = \frac{1}{100}, \beta = \frac{1}{100}, \omega = 52.5$$

ii.

$$\alpha = \frac{1}{2}, \beta = 1, \omega = \infty$$

iii.

$$\alpha = \frac{1}{2}, \beta = 2, \omega = -2.5$$

iv.

$$\alpha = 1, \beta = 1, \omega = 3$$

3. (a)

$$f(x) = \frac{1}{1+exp(-x)}, f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{t}{(1+t)^2} = \frac{t}{t^2+1+2t} = \frac{1}{t+\frac{1}{t}+2} \leq \frac{1}{2+2} = \frac{1}{4}$$

, where $t = e^{-x} \in (0, \infty)$, so Lipschitz constant of logistic function is $\frac{1}{4}$.

(b) $h(x) = g_1(g_2(x)), h'(x) = g_1'(g_2(x))g_2'(x) \leq C_1C_2$. Lipschitz constant is C_1C_2.

(c) $X_{j+1,i}(x) = f(a_{j+1,i}^T X_j(x)), X_{j+1,i}'(x) f'(a_{j+1,i}^T X_j(x))(a_{j+1,i}^T)X_{j,i}'(x) \leq \frac{1}{4}||a_{j+1,i}C_j|| \leq \frac{1}{4}||a_{j+1,i}||_\infty ||C_j||_\infty$. Then the Lipschitz constant of $X_{j+1,i}(x)$ is $\frac{1}{4}||a_{j+1,i}||_\infty ||C_j||_\infty$.

(d) From (c) we know that

$$X_{m,1}'(x) \leq \frac{1}{4}||a_{m,1}||_\infty \cdot X_{m-1,1}'(x) \leq \frac{1}{4}a \cdot (X_{m-1,1}'(x)) \leq \frac{1}{4}a(\frac{1}{4}a \cdot X_{m-2,1}'(x)) = (\frac{1}{4}a)^2 X_{m-2,1}'(x)$$

$$\leq ... \leq (\frac{1}{4}a)^m X_{m-m,1}'(x) = (\frac{1}{4}a)^m X_{0,1}'(x) = (\frac{1}{4}a)^m x_i'(x) = (\frac{1}{4}a)^m.$$

4. (a) Höffding's inequality cannot be applied directly here because the first term is $sup\hat{E}_n[f(X)]$ instead of a sum term, but Theorem 6.3.1 under this section can be applied. According to the Theorem 6.3.1,

$$\mathbb{P}(\sup_{g\in\mathcal{G}_n} |\hat{E}_n[g(X)] - E[g(X)]| \geq \epsilon) \leq 2\mathcal{N}(\frac{\epsilon}{3}, \mathcal{G}_n, ||\cdot||_\infty)exp(-2n(\frac{\epsilon}{3B})^2)$$

, for $\mathcal{G}_n$ a set of functions $g: \mathbb{R}^d \to [0, B]$. $B = 1$ here, and just taking one side of this formula:

$$\mathbb{P}(\sup_{f\in\mathcal{F}_C} \hat{E}_n[f(X)] - E[f(X)] \geq \epsilon) \leq \mathcal{N}(\frac{\epsilon}{3}, \mathcal{F}_C, ||\cdot||_\infty)exp(-2C(\frac{\epsilon}{3})^2)$$

$$\leq (\frac{8}{\frac{\epsilon}{3}})^{((\frac{6C}{\epsilon})^d)} exp(-2C(\frac{\epsilon}{3})^2) = (\frac{24}{\epsilon})^{((\frac{6C}{\epsilon})^d)} exp(-\frac{2}{9}\epsilon^2 C).$$

(b) Using Sauer's Lemma:

$$\mathbb{P}(\sup_{f\in\mathcal{F}} |\hat{E}_n[L(Y,f)] - E[L(Y,f)]| \geq \epsilon) \leq 2(\frac{en}{d})^d exp(-2n\epsilon^2)),$$

it can be shown that

$$\mathbb{P}(\sup_{f\in\mathcal{F}} \hat{E}_n[L(Y,f)] - E[L(Y,f)] \geq \epsilon) \leq (\frac{en}{d})^d exp(-2n\epsilon^2)),$$

then

$$\mathbb{P}(\hat{E}_n[L(Y,f)] - E[L(Y,f)] \geq \epsilon) \leq \mathbb{P}(\sup_{f\in\mathcal{F}} \hat{E}_n[L(Y,f)] - E[L(Y,f)] \geq \epsilon) \leq (\frac{en}{d})^d exp(-2n\epsilon^2)),$$

i.e.

$$\mathbb{P}(E_n[L(Y,\hat{f})] - E[L(Y,f^*)] \geq \epsilon) \leq (\frac{en}{d})^d exp(-2n\epsilon^2))$$

(c) Substitute δ with $(1-\delta)$ to range n in Sauer's Lemma:

$$n \leq \frac{128dlog(\frac{64d}{\epsilon^2}) + 64dlog(\frac{e}{d} + 16log(\frac{4}{1-\delta}))}{\epsilon^2}$$

$$n \geq \frac{max(\frac{1}{2}log(\frac{1}{4(1-\delta)}), 8d)}{\epsilon^2}$$