HW 2 Solution

Before starting let's prove the decomposition of the error in estimation an approximation terms. Let $\langle f,g\rangle = E[f(x)g(x)]$, and notice this is an inner product. Therefore, by the Cauchy-Schwarz (CS) inequality $|\langle f,g\rangle|^2 \leq \langle f,f\rangle\langle g,g\rangle \Rightarrow \langle f,g\rangle \leq \sqrt{E[f^2(x)]E[g^2(x)]}$

Now

$$
\begin{aligned}
E\left[(\hat{f}_n(x)-f^*(x))^2\right] &= E\left[(\hat{f}_n(x)-\bar{f}(x)+\bar{f}(x)-f^*(x))^2\right] \\
&= E\left[(\hat{f}_n(x)-\bar{f}(x))^2\right] + E\left[(\bar{f}(x)-f^*(x))^2\right] \\
&\quad + 2E\left[(\hat{f}_n(x)-\bar{f}(x))(\bar{f}(x)-f^*(x))\right] \\
\text{(by CS ineq)} \quad &\leq E\left[(\hat{f}_n(x)-\bar{f}(x))^2\right] + E\left[(\bar{f}(x)-f^*(x))^2\right] \\
&\quad + 2\sqrt{E\left[(f_n(x)-\bar{f}(x))^2\right]E\left[(\bar{f}(x)-f^*(x))^2\right]}
\end{aligned}
$$

Now let's study $E\left[(\hat{f}_n(x)-\bar{f}(x))^2\right]$ and $E\left[(\bar{f}(x)-f^*(x))^2\right]$ for the choice of $\bar{f}$ in the problem.

1 -

$$
\begin{aligned}
E\left[(\bar{f}(x)-f^*(x))^2\right] &= \int_{[0,1]} (\bar{f}(x)-f^*(x))^2\, dP_X(x) \\
&= \sum_{j=1}^{m} \int_{I_j} (\bar{c}_j - f^*(x))^2\, dP_X(x) \\
&= \sum_{j:\int_{I_j} dP_X(y)\neq 0} \int_{I_j} \left(\frac{\int_{I_j} f^*(y)\,dP_X(y)}{\int_{I_j} dP_X(y)} - f^*(x)\right)^2 dP_X(x) \\
&\quad + \underbrace{\sum_{j:\int_{I_j} dP_X(y)=0} \int_{I_j} (f^*(x))^2\, dP_X(x)}_{=0}
\end{aligned}
$$

$$= \sum_{j : \int_{I_j} dP_X(y) \neq 0} \int_{I_j} \left(\frac{\int_{I_j} (f^*(y) - f^*(x)) \, dP_X(y)}{\int_{I_j} dP_X(y)} \right)^2 dP_X(x) + 0$$

$$\leq \sum_{j : \int_{I_j} dP_X(y) \neq 0} \int_{I_j} \left(\frac{\int_{I_j} |f^*(y) - f^*(x)| \, dP_X(y)}{\int_{I_j} dP_X(y)} \right)^2 dP_X(x)$$

$$\leq \sum_{j : \int_{I_j} dP_X(y) \neq 0} \int_{I_j} \left(\frac{\int_{I_j} \frac{L}{m} \, dP_X(y)}{\int_{I_j} dP_X(y)} \right)^2 dP_X(x)$$

$$\leq \sum_{j=1}^{m} \int_{I_j} \left(\frac{L}{m}\right)^2 dP_X(x) = \left(\frac{L}{m}\right)^2 = \mathcal{O}\left(\frac{1}{m^2}\right) \quad \text{as } m \to \infty$$

2 - Let $N_j = \sum_{i=1}^{n} 1\{X_i \in I_j\}$ and $B_j := \sum_{i=1}^{n} Y_i \, 1\{X_i \in I_j\}$

Note that $N_j \sim \mathrm{Bin}(n, p_j)$, where $p_j = P(X \in I_j)$

Let's look first at $E\left[(\hat{f}(x) - \bar{f}(x))^2\right]$ for a fixed $x \in [0,1]$.
Let x be such that $x \in I_j$. Convention $\frac{0}{0} = 0$, so that if $N_j = 0$ $\frac{B_j}{N_j} = 0$. Then

$$E\left[\left(\hat{f}(x)-\bar{f}(x)\right)^2\right] = E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right)^2\right].$$

Notice that conditionally on $N_j = k$, B_j is the sum of k iid r.v.'s with mean $\bar{c}_j$. Also since $Y_i \in [-r, r]$ the variance of Y_i is at most r^2, therefore ~~Bj is the sum~~ given $N_j = k$ B_j is the sum of k iid r.v.'s with mean $\bar{c}_j$ and variance $\leq r^2$. So

$$E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right)^2 \Big| N_j = k\right] = \begin{cases} E\left[\left(\frac{B_j}{k}-\bar{c}_j\right)^2 \Big| N_j = k\right] & , k \neq 0 \\ E\left[(\bar{c}_j)^2 \mid N_j = 0\right] & , k = 0 \end{cases}$$

If $N_j \neq k$ $\quad E\left[\frac{(B_j - k\bar{c}_j)^2}{k^2} \Big| N_j = k\right] \leq \frac{r^2}{k}$

If $N_j = 0$ $\quad E\left[(\bar{c}_j)^2 \mid N_j = 0\right] \leq r^2$

Therefore $\quad E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right)^2\right] = E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right) 1\{N_j > 0\} + \left(\frac{B_j}{N_j}-\bar{c}_j\right) 1\{N_j = 0\}\right]$

$$\leq E\left[\frac{r^2}{N_j} 1\{N_j > 0\}\right] + r^2 \underbrace{E\left[1\{N_j = 0\}\right]}_{P(N_j = 0) = (1-p_j)^n}$$

We can now use the fact about binomials and get

$$E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right)^2\right] \leq \begin{cases} r^2 \frac{2}{(n+1)p_j} + r^2(1-p_j)^n & \text{if } p_j \neq 0 \\ r^2 & \text{if } p_j = 0 \end{cases}$$

We are almost done. We need to bound $E\left[(\hat{f}_n(X)-\bar{f}(X))^2\right]$ and

so

$$E\left[(\hat{f}_n(X)-\bar{f}(X))^2\right] = E\left[\sum_{j=1}^{m} (\hat{f}_n(X)-\bar{f}(X))^2 1\{X\in I_j\}\right]$$

$$= \sum_{j=1}^{m} E\left[\left(\frac{B_j}{N_j}-\bar{c}_j\right)^2 1\{X\in I_j\}\right]$$

$$\leq \sum_{j=1}^{m} E\left[\left(r^2\frac{2}{(n+1)p_j}+r^2(1-p_j)^n\right)1\{X\in I_j\}\right]$$

$$= \sum_{j=1}^{m}\left(r^2\frac{2}{(n+1)p_j}+r^2(1-p_j)^n\right)P(X\in I_j)$$

$$= \sum_{j=1}^{m}\left(r^2\frac{2}{(n+1)p_j}+r^2(1-p_j)^n\right)p_j$$

$$= \frac{2r^2m}{n+1} + r^2\sum_{j=1}^{m}(1-p_j)^n p_j \leq \frac{2r^2m}{n+1} + r^2 m\frac{1}{n+1} = \theta\left(\frac{m}{n+1}\right) = \theta\left(\frac{m}{n}\right)$$

as $\frac{m}{n}\to\infty$

The final step follows from the fact that

$$(1-p)^n p \leq \left(1-\frac{1}{n+1}\right)^n\frac{1}{n+1} \leq \left(e^{-\frac{1}{n+1}}\right)^n\frac{1}{n+1} \leq \frac{1}{n+1}$$

since $1-x\leq e^{-x}$

3 We have all the pieces now and

$$E\left[\left(\hat{f}_n(X) - f^*(X)\right)^2\right] = \mathcal{O}\left(\max\left\{\frac{1}{m^2}, \frac{m}{n}, \sqrt{\frac{1}{m^2}\frac{m}{n}}\right\}\right)$$

therefore taking $m \sim n^{1/3}$ yields

$$E\left[\left(\hat{f}_n(X) - f^*(X)\right)^2\right] = \mathcal{O}\left(\max\left\{n^{-\frac{2}{3}}, n^{-2/3}, \sqrt{n^{-2/3} \cdot n^{-2/3}}\right\}\right)$$

$$= \mathcal{O}\left(n^{-2/3}\right).$$

This is the same rate we had before, for the fixed design approach.