(6) Consider nonparametric regression with

$$Y_i = f(x_i) + \epsilon_i, \qquad i = 1, \ldots, n,$$

with the ϵ_i uncorrelated with common variance σ^2, and $x_i = i/n$, $i = 1, \ldots$ (so the x_i are evenly spaced). Assume that f is twice continuously differentiable. Consider estimating f by averaging Y_i with x_i near x. Specifically, let c be a positive constant, take

$$m = m(x) = \#\{i : |x_i - x| < cn^{-1/5}\},$$

so $m \sim 2cn^{4/5}$ as $n \to \infty$ for $x \in (0, 1)$, and define

$$\hat{f}(x) = \frac{1}{m} \sum_{i:|x_i - x| < cn^{-1/5}} Y_i.$$

(a) Derive an approximation for the bias of $\hat{f}(1/2)$ as $n \to \infty$ with an approximation error which is $o(n^{-2/5})$.

(b) Derive an approximation for the mean square error for $f(1/2)$ with error $o(n^{-4/5})$. What choice for c will make this mean square error as small as possible for large n? (The choice will depend on derivatives of f at $1/2$.)