Solutions to selected exercises from Chapter 20 of *Wasserman — All of Statistics*

(1) (a) We have

$$\mathbb{E}(\hat{f}_n(x)) = \frac{1}{nh} \sum \mathbb{E}\left(K\left(\frac{x - X_i}{h}\right)\right)$$

and

$$K\left(\frac{x - X_i}{h}\right) = \begin{cases} 1 & |x - X_i| \leq h/2 \\ 0 & \text{otherwise.} \end{cases}$$

Hence the expected value of $K((x - X_i)/h)$ is

$$\int K\left(\frac{x - y}{h}\right) f(y)dy = \int_{x-h/2}^{x+h/2} 1 \cdot f(y)dy.$$

Plugging this in to the sum for $\mathbb{E}(\hat{f}_n(x))$ gives the identity for $\mathbb{E}(\hat{f}_n(x))$.

By independence of the random variables $K((x - X_i)/h)$, we have

$$\mathbb{V}(\hat{f}_n(x)) = \frac{1}{n^2h^2} \sum \mathbb{V}\left(K\left(\frac{x - X_i}{h}\right)\right).$$

We have

$$\mathbb{V}\left(K\left(\frac{x - X_i}{h}\right)\right) = \mathbb{E}\left(K\left(\frac{x - X_i}{h}\right)^2\right) - \mathbb{E}\left(K\left(\frac{x - X_i}{h}\right)\right)^2.$$

But $K((x - X_i)/h)^2 = K((x - X_i)/h)$ so by our previous calculations,

$$\mathbb{V}\left(K\left(\frac{x - X_i}{h}\right)\right) = \int_{x-h/2}^{x+h/2} f(y)dy - \left(\int_{x-h/2}^{x+h/2} f(y)dy\right)^2$$

plugging this in to the sum for $\mathbb{V}(\hat{f}_n(x))$ completes the proof.

(b) We have that

$$\frac{1}{h}\int_{x-(h/2)}^{x+(h/2)} f(y)dy$$

converges to the derivative of $F(y) = \int_0^y f(y)dy$ at x, i.e. $f(x)$.

We may write the variance as

$$\frac{1}{nh} \cdot \frac{1}{h}\int_{x-(h/2)}^{x+(h/2)} f(y)dy - \frac{1}{n} \cdot \frac{1}{h^2}\left(\int_{x-(h/2)}^{x+(h/2)} f(y)dy\right)^2.$$

By our previous calculations $(\int_{x-(h/2)}^{x+(h/2)} f(y)dy)/h$ and $(\int_{x-(h/2)}^{x+(h/2)} f(y)dy)^2/h^2$ converge to $f(x)$ and $f(x)^2$, respectively. Since $1/nh \to 0$ and $1/n \to 0$ this implies that $\mathbb{V}(\hat{f}_n(x)) \to 0$.

Thus, $\hat{f}_n(x)$ converges to $f(x)$ in quadratic mean and therefore also in probability.

(2) See the Jupyter Notebook 2.ipynb.

(3) See the Jupyter Notebook 3.ipynb.

(4) We have

$$\mathbb{E}(L(g,\hat{g})) = \mathbb{E}\left(\int (gx - \hat{g}x)^2 dx\right) = \int \mathbb{E}(gx - \hat{g}x)^2 dx$$

by changing the order of integration. This is equal to

$$\int \mathbb{E}(gx - \mathbb{E}(\hat{g}x) + \mathbb{E}(\hat{g}x) - \hat{g}x)^2 dx$$

The integrand is equal to

$$v(x) + \mathbb{E}((gx - \mathbb{E}(\hat{g}x))^2 - 2(gx - \mathbb{E}(\hat{g}x))(\hat{g}x - \mathbb{E}(\hat{g}x)))$$

so we need to confirm

$$\mathbb{E}((gx - \mathbb{E}(\hat{g}x))^2 - 2(gx - \mathbb{E}(\hat{g}x))(\hat{g}x - \mathbb{E}(\hat{g}x))) = b(x)^2.$$

We expand the argument to $\mathbb{E}(\cdot)$ as

$$gx^2 - 2gx\mathbb{E}(\hat{g}x) + (\mathbb{E}\hat{g}x)^2 - 2(gx)(\hat{g}x) + 2(gx)\mathbb{E}(\hat{g}x) + 2(\hat{g}x)\mathbb{E}(\hat{g}x) - 2(\mathbb{E}\hat{g}x)^2$$

and this is equal to

$$-(\mathbb{E}\hat{g}x)^2 + gx^2 - 2(gx)(\hat{g}x) + 2(\hat{g}x)\mathbb{E}(\hat{g}x).$$

Taking the expectation yields

$$-(\mathbb{E}\hat{g}x)^2 + gx^2 - 2gx(\mathbb{E}\hat{g}x) + 2(\mathbb{E}\hat{g}x)^2 = (\mathbb{E}\hat{g}x)^2 - 2(gx)(\mathbb{E}\hat{g}x) + (gx)^2 = b(x)^2$$

as claimed.

(5) The derivation $\mathbb{E}(\hat{f}_n(x)) = p_j/h$ was given on page 306. For the expression for $\mathbb{V}(\hat{f}_n(x))$, we have

$$\mathbb{V}(\hat{f}_n(x)) = \mathbb{V}(\nu_j/nh) = \frac{1}{n^2h^2}\mathbb{V}\left(\sum_{i=1}^n I_{B_j}(X_i)\right) = \frac{1}{n^2h^2}\sum_{i=1}^n \mathbb{V}(I_{B_j}(X_i))$$

by independence of the random variables $I_{B_j}(X_i)$ (where I_{B_j} denotes the indicator function for B_j). Since $I_{B_j}^2 = I_{B_j}$, we have

$$\mathbb{V}(I_{B_j}(X_i)) = \mathbb{E}(I_{B_j}(X_i)) - \mathbb{E}(I_{B_j}(X_i))^2 = p_j - p_j^2.$$

Plugging this in yields

$$\mathbb{V}(\hat{f}_n(x)) = n\frac{1}{n^2h^2}p_j(1-p_j),$$

as claimed.

(6) First consider the term $\int \hat{f}_n dx$. This is equal to

$$\sum_j \int_{B_j} \hat{f}_n^2 dx = \sum_j \int_{B_j} \frac{\hat{p}_j^2}{h^2} dx = \frac{1}{h}\sum_j \hat{p}_j^2.$$

Now consider the term $\frac{2}{n}\sum_i \hat{f}_{(-i)}(X_i)$. Suppose $X_i \in B_j$. Then $\hat{f}_{(-i)}(X_i) = \frac{\nu_j - 1}{(n-1)h}$ since, leaving out X_i, there are $\nu_j - 1$ X_k's in B_j and $n-1$ X_k's total. Breaking the sum over i into a sum over the X_i's in B_j for each j yields

$$\sum_i \hat{f}_{(-i)}(X_i) = \sum_j \nu_j \left(\frac{\nu_j - 1}{(n-1)h} \right) = \sum_j \frac{\nu_j^2}{(n-1)h} - \sum_j \frac{\nu_j}{(n-1)h}$$

since there are ν_j X_i's in B_j for each j. Finally then,

$$\sum_i \hat{f}_{(-i)}(X_i) = \frac{n^2}{(n-1)h} \sum \hat{p}_j^2 - \frac{n}{(n-1)h}.$$

So the cross-validation estimator of risk is

$$\frac{1}{h}\sum_j \hat{p}_j^2 - \frac{2n}{(n-1)h}\sum \hat{p}_j^2 + \frac{2}{(n-1)h} = \frac{2}{(n-1)h} - \frac{n+1}{(n-1)h}\sum \hat{p}_j^2,$$

as claimed.