Ex. 6.8

Some notation to be explicit:

$$K_\lambda(x_0, x) = \phi\left(\frac{|x - x_0|}{\lambda}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{|x - x_0|}{\lambda}\right)^2\right)$$

$$= \phi_\lambda(x - x_0)$$

So $\hat{f}_X(x) = \frac{1}{N\lambda} \sum_{i=1}^{N} \phi_\lambda(x - x_i)$

also $\hat{f}_{X,Y}(x,y) = \frac{1}{N\lambda^2} \sum_{i=1}^{N} \phi_\lambda(x - x_i)\,\phi_\lambda(y - y_i)$

Now by definition of conditional expectation:

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y\, \hat{f}_{Y|X}(y \mid x)\, dy$$

$$= \int_{-\infty}^{\infty} y \left(\frac{\hat{f}_{X,Y}(x,y)}{\hat{f}_X(x)}\right) dy$$

$$= \int_{-\infty}^{\infty} \frac{(1/N\lambda^2)\, y \sum_{i=1}^{N} \phi_\lambda(x - x_i)\,\phi_\lambda(y - y_i)}{(1/N\lambda) \sum_{i=1}^{N} \phi_\lambda(x - x_i)}\, dy \qquad \text{(subbing in)}$$

$$= \frac{\sum_{i=1}^{N} \left[\phi_\lambda(x - x_i) \int_{-\infty}^{\infty} y\, \phi_\lambda(y - y_i)\, dy\right]}{\lambda \sum_{i=1}^{N} \phi_\lambda(x - x_i)} \qquad (*)$$

Now, for a constant K, we can show:

$$\int_{-\infty}^{\infty} y\, \phi_\lambda(y - K)\, dy = \lambda K \qquad \text{for } K \geq 0$$

Then $(*)$ becomes: $\dfrac{\sum_{i=1}^{N} \phi_\lambda(x - x_i)\, y_i}{\sum_{i=1}^{N} \phi_\lambda(x - x_i)}$ which is a Nadaraya-Watson estimator as required.

For the classification setting we can simply replace $\phi_\lambda(y-y_i)$ with $\delta(y_i, g)$ where:

$$\delta(y_i, g) = \begin{cases} 1 & \text{if } y_i = g \\ 0 & \text{otherwise} \end{cases}$$

Now,

$$\mathbb{P}(G=g \mid X=x) := \frac{\hat{f}_{XG}(x, g)}{\hat{f}_X(x)}$$

$$= \frac{(1/N\lambda)\sum_{i=1}^{N} \phi_\lambda(x-x_i)\,\delta(y_i, g)}{(1/N\lambda)\sum_{i=1}^{N} \phi_\lambda(x-x_i)}$$

$$= \frac{\sum_{g_i=g} \phi_\lambda(x-x_i)}{\sum_{i=1}^{N} \phi_\lambda(x-x_i)}$$