Q1:

Sample points are represented by "y".

$$E\left(\hat{f}(x)\right) = E\left[\frac{1}{nh}\sum_y K\left(\frac{x-y}{h}\right)\right] = \frac{1}{nh}\sum_y E\left[K\left(\frac{x-y}{h}\right)\right] = \frac{1}{h}E[K\left(\frac{x-y}{h}\right)] = \frac{1}{h}\int_{-\infty}^{+\infty} K\left(\frac{x-y}{h}\right)f(y)dy =$$

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

Using formula 24, page 193 of "statistical classification" book written by Duda et al.:

$$V\left(\hat{f}(x)\right) = \frac{1}{nh^2}\int_{-\infty}^{+\infty} K^2\left(\frac{x-y}{h}\right)f(y)dy - \frac{1}{n}E^2\left(\hat{f}(x)\right) = \frac{1}{nh^2}\int_{y=x-\left(\frac{h}{2}\right)}^{y=x+\left(\frac{h}{2}\right)} 1 \times f(y)dy - \frac{1}{n}\left(\frac{1}{h}\int_{x-\left(\frac{h}{2}\right)}^{x+\left(\frac{h}{2}\right)} f(y)dy\right)^2$$

$$= \frac{1}{nh^2}\left[\int_{y=x-\left(\frac{h}{2}\right)}^{y=x+\left(\frac{h}{2}\right)} f(y)dy - \left(\int_{x-\left(\frac{h}{2}\right)}^{x+\left(\frac{h}{2}\right)} f(y)dy\right)^2\right]$$

$$V\left(\hat{f}(x)\right) = V\left[\frac{1}{nh}\sum_y K\left(\frac{x-y}{h}\right)\right] = \sum_y V\left[\frac{1}{nh}K\left(\frac{x-y}{h}\right)\right] = nV\left[\frac{1}{nh}K\left(\frac{x-y}{h}\right)\right]$$

$$= n\left[E\left[\frac{1}{n^2h^2}K^2\left(\frac{x-y}{h}\right)\right] - E^2\left[\frac{1}{nh}K\left(\frac{x-y}{h}\right)\right]\right]$$

$$= n\left[\frac{1}{n^2h^2}\int_{-\infty}^{+\infty} K^2\left(\frac{x-y}{h}\right)f(y)dy - \frac{1}{h^2}\left(\int_{x-\left(\frac{h}{2}\right)}^{x+\left(\frac{h}{2}\right)} f(y)dy\right)^2\right]$$

$$= \frac{1}{h^2}\left[\frac{1}{n}\int_{y=x-\left(\frac{h}{2}\right)}^{y=x+\left(\frac{h}{2}\right)} f(y)dy - n\left(\int_{x-\left(\frac{h}{2}\right)}^{x+\left(\frac{h}{2}\right)} f(y)dy\right)^2\right]$$

Q2:

Useful formula: $\int_{-\infty}^{\infty} \exp(-a(x+b)^2)\,dx = \sqrt{\frac{\pi}{a}}$

Part A)

$$E(\hat{p}_N(x)) = \int \frac{1}{h}\Phi\left(\frac{x-v}{h}\right)p(v)dv$$

$$= \int \frac{1}{h\sqrt{2\pi}}\exp\left(-\frac{(x-v)^2}{2h^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{(v-\mu)^2}{\sigma^2}\right)dv = \frac{1}{2\pi h\sigma}\int \exp\left(-\frac{1}{2}\left[\frac{(v-x)^2}{h^2}+\frac{(v-\mu)^2}{\sigma^2}\right]\right)dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp\left(-\frac{1}{2}\left[\frac{\sigma^2(v-x)^2+h^2(v-\mu)^2}{h^2\sigma^2}\right]\right)dv =$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{\sigma^2 v^2+\sigma^2 x^2-2\sigma^2 vx+h^2 v^2+h^2\mu^2-2h^2 v\mu}{h^2\sigma^2}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{v^2(\sigma^2+h^2)-2v(\sigma^2 x+h^2\mu)+h^2\mu^2+\sigma^2 x^2}{h^2\sigma^2}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{v^2-\frac{2v(\sigma^2 x+h^2\mu)}{(\sigma^2+h^2)}+\frac{h^2\mu^2}{(\sigma^2+h^2)}+\frac{\sigma^2 x^2}{(\sigma^2+h^2)}}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}])\,dv =$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{v^2-\frac{2v(\sigma^2 x+h^2\mu)}{(\sigma^2+h^2)}+\frac{h^2\mu^2}{(\sigma^2+h^2)}+\frac{\sigma^2 x^2}{(\sigma^2+h^2)}+\frac{(\sigma^2 x+h^2\mu)^2}{(\sigma^2+h^2)^2}-\frac{(\sigma^2 x+h^2\mu)^2}{(\sigma^2+h^2)^2}}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{(v-\frac{h^2\mu+\sigma^2 x}{h^2+\sigma^2})^2+\frac{h^2\mu^2}{(\sigma^2+h^2)}+\frac{\sigma^2 x^2}{(\sigma^2+h^2)}-\frac{(\sigma^2 x+h^2\mu)^2}{(\sigma^2+h^2)^2}}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{(v-\frac{h^2\mu+\sigma^2 x}{h^2+\sigma^2})^2+\frac{\sigma^2 h^2\mu^2+h^4\mu^2+\sigma^4 x^2+\sigma^2 h^2\mu^2-\sigma^4 x^2-h^4\mu^2+2\sigma^2 xh^2\mu}{(\sigma^2+h^2)^2}}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{(v-\frac{h^2\mu+\sigma^2 x}{h^2+\sigma^2})^2+\frac{\sigma^2 h^2(x-\mu)^2}{(\sigma^2+h^2)^2}}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}])\,dv = \frac{1}{2\pi h\sigma}\int \exp(-\frac{1}{2}[\frac{(v-\frac{h^2\mu+\sigma^2 x}{h^2+\sigma^2})^2}{\frac{h^2\sigma^2}{(\sigma^2+h^2)}}+\frac{(x-\mu)^2}{(\sigma^2+h^2)}])\,dv$$

$$= \frac{1}{2\pi h\sigma}\int_{-\infty}^{\infty} \exp(-\frac{1}{2}\frac{(v-\frac{h^2\mu+\sigma^2 x}{h^2+\sigma^2})^2}{\frac{h^2\sigma^2}{h^2+\sigma^2}})\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{h^2+\sigma^2}\right)dv = \frac{\sqrt{2\pi}}{2\pi h\sigma}\left(\frac{h\sigma}{\sqrt{h^2+\sigma^2}}\right)\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2+h^2}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+h^2}}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2+h^2}\right) = N(\mu,\sigma^2+h^2)$$

It has a simpler solution if we use the properties of the convolution of two Gaussians.

Part B)

From above, we know: $\int \exp\left(-\frac{1}{2}\left[\frac{(v-x)^2}{h^2} + \frac{(v-\mu)^2}{\sigma^2}\right]\right) dv = \sqrt{2\pi}\left(\frac{h\sigma}{\sqrt{h^2+\sigma^2}}\right)\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2+h^2}\right)$

Another useful formula: $N^2(a,b^2) = \left(\frac{1}{\sqrt{2\pi}b}\exp(-\frac{(x-a)^2}{2b^2})\right)^2 = \frac{1}{\sqrt{2\pi}b\sqrt{2\pi}b}\exp\left(-\frac{(x-a)^2}{2\frac{b^2}{2}}\right) = \frac{1}{2b\sqrt{\pi}}N(a,\frac{b^2}{2})$

$$Var(\hat{p}_N(x)) = \frac{1}{nh^2}\int \Phi^2\left(\frac{x-v}{h}\right)p(v)dv - \frac{1}{n}E^2(\hat{p}_N(x))$$

$$= \frac{1}{nh^2}\int\left[\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(v-x)^2}{h^2}\right)\right]^2 N(\mu,\sigma^2)dv - \frac{1}{n}N^2(\mu,\sigma^2+h^2)$$

$$= \frac{1}{nh^2}\int\left[\frac{1}{2\pi}\exp\left(-\frac{1}{2}\frac{(v-x)^2}{\frac{h^2}{2}}\right)\right]N(\mu,\sigma^2)dv - \frac{1}{n}N^2(\mu,\sigma^2+h^2)$$

$$= \frac{1}{nh^2}\int\left[\frac{h}{2\sqrt{\pi}}N\left(x,\frac{h^2}{2}\right)\right]N(\mu,\sigma^2)dv - \frac{1}{n}N^2(\mu,\sigma^2+h^2)$$

$$= \frac{1}{2\sqrt{\pi}nh}\times\frac{1}{h\sqrt{\pi}}\times\frac{1}{\sigma\sqrt{2\pi}}\int \exp\left(-\frac{1}{2}\left[\frac{(v-x)^2}{\frac{h^2}{2}} + \frac{(v-\mu)^2}{\sigma^2}\right]\right)dv - \frac{1}{n}N^2(\mu,\sigma^2+h^2) =$$

$$= \frac{1}{2\sqrt{\pi}nh}\times\frac{1}{h\sqrt{\pi}}\times\frac{1}{\sigma\sqrt{2\pi}}\left(\frac{\sqrt{2\pi}\frac{h}{\sqrt{2}}\sigma}{\sqrt{\frac{h^2}{2}+\sigma^2}}\right)\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2+\frac{h^2}{2}}\right) - \frac{N\left(\mu,\frac{\sigma^2+h^2}{2}\right)}{2\sqrt{\pi}n\sqrt{\sigma^2+h^2}}$$

$$= \frac{N\left(\mu,\sigma^2+\frac{h^2}{2}\right)}{2\sqrt{\pi}nh} - \frac{N\left(\mu,\frac{\sigma^2+h^2}{2}\right)}{2\sqrt{\pi}n\sqrt{\sigma^2+h^2}} = \frac{1}{n}\left(\frac{\textcolor{red}{N\left(\mu,\sigma^2+\frac{h^2}{2}\right)}}{\textcolor{red}{2\sqrt{\pi}h}} - \frac{\textcolor{blue}{N\left(\mu,\frac{\sigma^2+h^2}{2}\right)}}{\textcolor{blue}{2\sqrt{\pi}\sqrt{\sigma^2+h^2}}}\right) \cong \frac{N(\mu,\sigma^2)}{2\sqrt{\pi}nh}$$

The red term has a large value for a small h. However, the blue term is much smaller and we can ignore it because $h^2 \cong 0$, and $\sqrt{\sigma^2+h^2} \cong \sigma$.

Q3-

Part A)

$$P_{N1}(x|\omega_1) = \frac{1}{N1} \sum_{i=1}^{N1} \frac{1}{h_{N1}} \Phi(\frac{x - x_{1i}}{h_{N1}})$$

$$P_{N2}(x|\omega_2) = \frac{1}{N2} \sum_{i=1}^{N2} \frac{1}{h_{N2}} \Phi(\frac{x - x_{2i}}{h_{N2}})$$

Part B) Bayes rule: $p(x|\omega_1)p(\omega_1) \gtrless_{\omega_2}^{\omega_1} p(x|\omega_2)p(\omega_2)$ and from the question: $p(\omega_1) = p(\omega_2)$

$$\frac{1}{N1} \sum_{i=1}^{N1} \frac{1}{h_{N1}} \Phi(\frac{x - x_{1i}}{h_{N1}}) \gtrless_{\omega_2}^{\omega_1} \frac{1}{N2} \sum_{i=1}^{N2} \frac{1}{h_{N2}} \Phi(\frac{x - x_{2i}}{h_{N2}})$$

and $h_{N1} = h_{N2}$

Part C)

$$\frac{1}{N1} \sum_{i=1}^{N1} x^T x_{1i} \gtrless_{\omega_2}^{\omega_1} \frac{1}{N2} \sum_{i=1}^{N2} x^T x_{2i} \rightarrow \frac{x^T}{N1} \sum_{i=1}^{N1} x_{1i} \gtrless_{\omega_2}^{\omega_1} \frac{x^T}{N2} \sum_{i=1}^{N2} x_{2i} \rightarrow x^T m_1 \gtrless_{\omega_2}^{\omega_1} x^T m_2 \rightarrow$$

This is nearest mean classifier. Why?

It is based on dot-product: $ab = |a||b|\cos\alpha \rightarrow$ This result is maximum when α is small and both a and b have similar length. So, the nearer to mean, the larger value for the probability.

Q4-

Part A) $v = h^d = h$ $E(\hat{p}_n) = \int \frac{1}{h} \Phi(\frac{x-v}{h})p(v)dv = \int \frac{1}{h} \exp(\frac{v-x}{h}) \frac{1}{a} dv = \frac{h}{ah} \exp(\frac{v-x}{h})|_?^?$
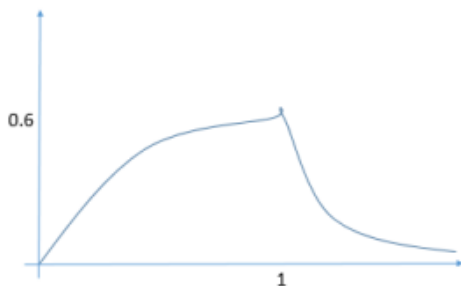
$x < 0$:   kernel is zero: 0

$0 < x < a$: $\frac{1}{a} \exp\left(\frac{v-x}{h}\right)|_0^x = \frac{1}{a} \exp(0) - \frac{1}{a} \exp\left(-\frac{x}{h}\right) = \frac{1}{a}(1 - \exp\left(-\frac{x}{h}\right))$
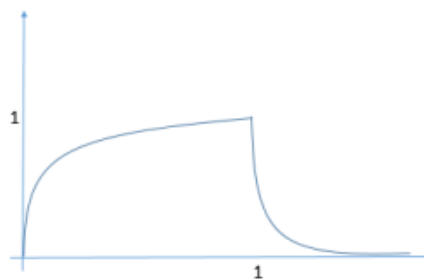
$x > a$: $p(v = 0)$, So: $\frac{1}{a} \exp\left(\frac{v-x}{h}\right)|_0^a = \frac{1}{a} \exp(-\frac{x}{h}) (\exp\left(\frac{a}{h}\right) - 1)$

Part B)

For $a = 1$ and $h = 1$:
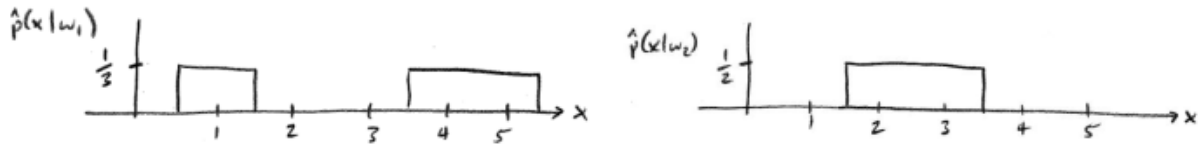

0.6

1

For $a = 1$ and $h = \frac{1}{4}$:

1

1

Part C) 99 percent of $[0\ a]$ is from 0 to $0.99a$. So, $x = 0.99a$:

$$\text{Bias} = E(p(x)) - E(\hat{p}_N(x)) = \frac{1}{a} - \frac{1}{a}\left(1 - e^{\frac{-x}{h}}\right) = \frac{1}{a}e^{\frac{-x}{h}} \qquad\qquad x = 0.01a$$

$$\frac{1}{a}e^{\frac{-0.99}{h}a} < \frac{1}{100}\left(\frac{1}{a}\right) \;\to\; \frac{-0.99a}{h} < -\ln(100) \;\to\; \frac{h}{0.99a} < \frac{1}{\ln(100)} \;\to\; h < \frac{0.99}{\ln(100)}a \;\to\; h < 0.215a$$

Q5:

Part A)

Part B) Posteriors are like likelihood except that the value on Y axis is 1 for them.

Q6:

Part A)

$$p(x) = \frac{1}{2 \times 2}\sum_{k=1}^{2} k\left(\frac{x - x_k}{2}\right) = \frac{1}{4}\left(k(0.3) + k(-0.2)\right) = \frac{1}{4}e^{-0.3} = 0.19$$
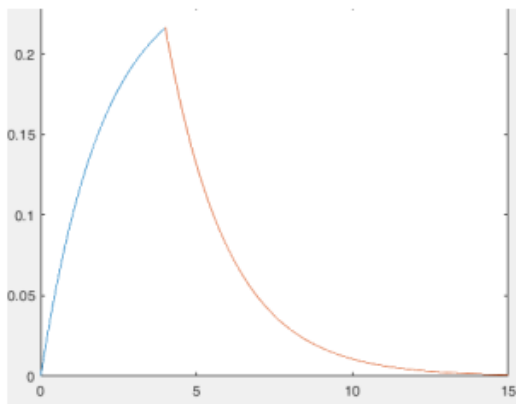
Part B)

$$E(\hat{p}_n) = \int \frac{1}{h}\Phi\left(\frac{x - v}{h}\right)p(v)dv = \int \frac{1}{2}\exp\left(\frac{v - x}{2}\right)\frac{1}{4}dv = \frac{2}{8}\exp\left(\frac{v - x}{2}\right)\Big|_?^?$$

$x < 0:\quad$ *kernel is zero*: 0

$0 < x < a:\quad \frac{1}{4}\left[\exp(0) - \exp\left(\frac{-x}{2}\right)\right] = \frac{1}{4}\left(1 - \exp\left(-\frac{x}{2}\right)\right)$

$x > a:\quad p(v = 0), So: \frac{1}{4}\left[\exp\left(\frac{4 - x}{2}\right) - \exp\left(\frac{-x}{2}\right)\right] = \frac{1}{4}\exp\left(-\frac{x}{2}\right)[\exp(2) - 1]$

In general, increasing h makes it smoother and less noisy but may ignore some details. However, decreasing h makes the pdf spiky and noisy.

Q8:

Part A)

$$p_n(x) = \frac{1}{nh}\sum_i K(u) \quad , \quad K(u) = \begin{cases} 1 & |u| \leq \frac{1}{2} \\ 0 & otherwise \end{cases}$$

$$|u| < \frac{1}{2} \quad \rightarrow \quad -\frac{1}{2} \leq u \leq \frac{1}{2} \quad \rightarrow \quad -\frac{1}{2} \leq \frac{x - x_i}{h} \leq \frac{1}{2} \quad \rightarrow \quad x_i - \frac{h}{2} \leq x \leq x_i + \frac{h}{2}$$

$$x_1 = 1: \quad 0 \leq x \leq 2$$

$$x_2 = 2: \quad 1 \leq x \leq 3$$

$$x_3 = 100: \quad 99 \leq x \leq 101$$

$$p(x) = \frac{1}{3 \times 2}\left(\begin{cases}1 & 0 \le x \le 2 \\ 0 & otherwise\end{cases} + \begin{cases}1 & 1 \le x \le 3 \\ 0 & otherwise\end{cases} + \begin{cases}1 & 99 \le x \le 101 \\ 0 & otherwise\end{cases}\right) = \begin{cases}\frac{1}{6} & 0 \le x \le 1 \\ \frac{2}{6} & 1 \le x \le 2 \\ \frac{1}{6} & 2 \le x \le 3 \\ 0 & 3 \le x \le 99 \\ \frac{1}{6} & 99 \le x \le 101 \\ 0 & otherwise\end{cases}$$

Part B)

Part C)

$$E(\hat{p}_n) = \int xp(x)dx = \int_0^1 \frac{1}{6}xdx + \int_1^2 \frac{2}{6}xdx + \int_2^3 \frac{1}{6}xdx + \int_{99}^{101} \frac{1}{6}xdx = \frac{1}{6}(\frac{1}{2}+3+\frac{5}{2}+200) = \frac{206}{6} = 34.33$$

Part D)

$$var(\hat{p}_n) = \int x^2 p(x)dx - E^2(\hat{p}_n) = \int_0^1 \frac{1}{6}x^2 dx + \int_1^2 \frac{2}{6}x^2 dx + \int_2^3 \frac{1}{6}x^2 dx + \int_{99}^{101} \frac{1}{6}x^2 dx - 34.33^2$$
$$= \frac{1}{6}(\frac{1}{3}+\frac{14}{3}+\frac{19}{3}+\frac{60002}{3}) - 34.33^2 = 3335.33 - 1178.55 = 2156.78$$

Q10-

$$k_n = \sum_{i=1}^n \varphi\left(\frac{\mathbf{x}-\mathbf{x}_i}{h_n}\right),$$

$$p_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n \frac{1}{V_n}\varphi\left(\frac{\mathbf{x}-\mathbf{x}_i}{h_n}\right).$$

Convergence of the Mean

$$\lim_{n \to \infty} \bar{p}_n(\mathbf{x}) = p(\mathbf{x})$$

$$
\begin{aligned}
\bar{p}_n(\mathbf{x}) &= \mathcal{E}[p_n(\mathbf{x})] \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathcal{E}\left[\frac{1}{V_n} \varphi\left(\frac{\mathbf{x} - \mathbf{x}_i}{h_n}\right)\right] \\
&= \int \frac{1}{V_n} \varphi\left(\frac{\mathbf{x} - \mathbf{v}}{h_n}\right) p(\mathbf{v}) \, d\mathbf{v} \\
&= \int \delta_n(\mathbf{x} - \mathbf{v}) p(\mathbf{v}) \, d\mathbf{v}.
\end{aligned}
$$

Convergence of the Variance

$$
\begin{aligned}
\sigma_n^2(\mathbf{x}) &= \sum_{i=1}^{n} \mathcal{E}\left[\left(\frac{1}{nV_n} \varphi\left(\frac{\mathbf{x} - \mathbf{x}_i}{h_n}\right) - \frac{1}{n}\bar{p}_n(\mathbf{x})\right)^2\right] \\
&= n\, \mathcal{E}\left[\frac{1}{n^2 V_n^2} \varphi^2\left(\frac{\mathbf{x} - \mathbf{x}_i}{h_n}\right)\right] - \frac{1}{n}\bar{p}_n^2(\mathbf{x}) \\
&= \frac{1}{nV_n} \int \frac{1}{V_n} \varphi^2\left(\frac{\mathbf{x} - \mathbf{v}}{h_n}\right) p(\mathbf{v}) \, d\mathbf{v} - \frac{1}{n}\bar{p}_n^2(\mathbf{x}).
\end{aligned}
$$

$$\sigma_n^2(\mathbf{x}) \leq \frac{\sup(\varphi(\cdot))\,\bar{p}_n(\mathbf{x})}{nV_n}.$$

راه دوم:

Now suppose $X_1, \ldots, X_n$ are i.i.d. with density f. Choose and fix a bandwidth $h > 0$ (small), and define

$$\hat{f}(x) := \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h}\right)$$

$$= \frac{1}{n} \sum_{j=1}^{n} K_h(x - X_j).$$

We can easily compute the mean and variance of $\hat{f}(x)$, viz.,

$$\mathrm{E}[\hat{f}(x)] = \mathrm{E}[K_h(x - X_1)]$$

$$= \int_{-\infty}^{\infty} K_h(x - y)f(y)\,dy = (K_h * f)(x);$$

$$\mathrm{Var}\,\hat{f}(x) = \frac{1}{n}\mathrm{Var}\,(K_h(x - X_1))$$

$$= \frac{1}{nh^2} \int_{-\infty}^{\infty} \left| K\left(\frac{x - y}{h}\right) \right|^2 f(y)\,dy - \frac{1}{n}|(K_h * f)(x)|^2$$

$$= \frac{1}{n}\left[(K_h^2 * f)(x) - (K_h * f)^2(x)\right],$$