Ex 7.2

We are told that $\mathbb{P}(y=1\mid x_0) = f(x_0)$ and
$G(x) = I(f(x) > 1/2)$

Now,

$$\text{Err}(x_0) = \mathbb{P}(Y \neq \hat{G}(x_0) \mid X = x_0)$$

There are two ways this can happen, $Y = G$ and $Y \neq G$. So:

$$= \mathbb{P}(Y = G(x_0) \mid X = x_0)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)$$
$$+ \mathbb{P}(Y \neq G(x_0) \mid X = x_0)\, \mathbb{P}(\hat{G}(x_0) = G(x_0) \mid X = x_0)$$

$$= \mathbb{P}(Y = G(x_0) \mid X = x_0)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)$$
$$+ \mathbb{P}(Y \neq G(x_0) \mid X = x_0)\left[1 - \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)\right]$$

From here there are two possible cases:

Case 1: $G(x_0) = 1$ e.g. $f(x_0) > 1/2$

$$\Rightarrow f(x_0)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)$$
$$+ (1 - f(x_0))\left[1 - \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)\right]$$

$$= (2f(x_0) - 1)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0) + 1 - f(x_0)$$

$$= (2f(x_0) - 1)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0) + \mathbb{P}(Y \neq G(x_0) \mid X = x_0)$$

Case 2: $G(x_0) = 0$ e.g. $f(x_0) \leq 1/2$

$$\Rightarrow (1 - f(x_0))\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)$$
$$+ f(x_0)\left[1 - \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)\right]$$

$$= -(2f(x_0) - 1)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x) + f(x_0)$$

$$= -(2f(x_0) - 1)\, \mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x) + \mathbb{P}(Y \neq G(x_0) \mid X = x_0)$$

Now, combining the results of Case 1 and Case 2 we can easily see that in general we have:

$$\text{Err}(x_0) = |2f(x_0)-1|\,\mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0) + \mathbb{P}(Y \neq G(x_0) \mid X = x_0)$$

$$= \text{Err}_B(x_0) + |2f(x_0)-1|\,\mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0)$$

(by definition of Err_B)

For the second part we will use the following fact:

If $X \sim N(\mu, \sigma^2)$ then $Z = \dfrac{X-\mu}{\sigma^2} \sim N(0,1)$

Using this fact we can again split into two cases:

Case 1: $f(x_0) > \frac{1}{2} \Rightarrow (\frac{1}{2} - f(x_0))$ is negative.

Then $\mathbb{P}(\hat{G}(x_0) \neq G(x_0)) = \mathbb{P}(\hat{G}(x_0) = 0)$

$$= \mathbb{P}(\hat{f}(x_0) < \tfrac{1}{2}) = \mathbb{P}\left(\frac{\hat{f}(x_0) - \mathbb{E}(\hat{f}(x_0))}{\sqrt{\text{Var}(\hat{f}(x_0))}} < \frac{\frac{1}{2} - \mathbb{E}(\hat{f}(x_0))}{\sqrt{\text{Var}(\hat{f}(x_0))}} \right)$$

Now for $Z \sim N(0,1)$, $\mathbb{P}(Z < x) = \Phi(x)$ so

$$= \Phi\left(\frac{\frac{1}{2} - \mathbb{E}(\hat{f}(x_0))}{\sqrt{\text{Var}(\hat{f}(x_0))}} \right)$$

Case 2: $f(x_0) < \frac{1}{2} \Rightarrow (\frac{1}{2} - f(x_0))$ is positive.

This case follows a similar logic with:

$$\mathbb{P}(\hat{G}(x_0) \neq G(x_0)) = \mathbb{P}(\hat{G}(x_0) = 1)$$

$$= \mathbb{P}(\hat{f}(x_0) > 1/2) = 1 - \mathbb{P}(\hat{f}(x_0) < 1/2)$$

$$= 1 - \Phi\left(\frac{1/2 - \mathbb{E}(\hat{f}(x_0))}{\sqrt{Var(\hat{f}(x_0))}}\right)$$ (Proved in case 1)

$$= \Phi\left(\frac{\mathbb{E}(\hat{f}(x_0)) - 1/2}{\sqrt{Var(\hat{f}(x_0))}}\right)$$ (Property of Φ Function)

Now it is clear that we can combine these two cases into a single expression by writing:

$$\mathbb{P}(\hat{G}(x_0) \neq G(x_0) \mid X = x_0) \approx \Phi\left(\frac{Sign(1/2 - f(x_0)[\mathbb{E}(\hat{f}(x_0) - 1/2]}{\sqrt{Var(\hat{f}(x_0))}}\right)$$

as required.