您提出的质疑非常正确，我在之前的证明中确实犯了一个错误。在步骤3中，我将比值的期望$E\left[\frac{A}{B}\right]$近似为期望的比值$\frac{E[A]}{E[B]}$，这是不严谨的。一般来说，$E\left[\frac{A}{B}\right] \neq \frac{E[A]}{E[B]}$。

请允许我重新进行证明，并详细解释每一个步骤。

重新证明：

目标：

证明对于任意给定的x，Nadaraya-Watson估计量$\hat{m}_h(x)$的无条件期望$E[\hat{m}_h(x)]$是否等于$E[Y]$，即验证$E[\hat{m}_h(x)] = E[Y]$是否成立。

步骤1：定义Nadaraya-Watson估计量

Nadaraya-Watson估计量用于估计回归函数$m(x) = E[Y|X = x]$，其形式为：

$$\hat{m}_h(x) = \frac{\sum_{i=1}^n K_h(x - X_i)Y_i}{\sum_{i=1}^n K_h(x - X_i)}$$

其中：

- $K_h(u) = \frac{1}{h}K\left(\frac{u}{h}\right)$，$K(\cdot)$是核函数；
- $h > 0$是带宽；
- $\{(X_i, Y_i)\}_{i=1}^n$是独立同分布的样本数据。

步骤2：考虑$\hat{m}_h(x)$的无条件期望$E[\hat{m}_h(x)]$

我们需要计算：

$$E[\hat{m}_h(x)] = E\left[\frac{\sum_{i=1}^n K_h(x - X_i)Y_i}{\sum_{i=1}^n K_h(x - X_i)}\right]$$

直接计算这个期望是困难的，因为期望的比值不等于比值的期望。

步骤3：引入条件期望

为了更好地处理这个问题，我们可以考虑对分母条件于$\{X_i\}$，然后计算条件期望。

定义：

$$S_n(x) = \sum_{i=1}^n K_h(x - X_i)$$

$$N_n(x) = \sum_{i=1}^{n} K_h(x - X_i)Y_i$$

因此，$\hat{m}_h(x) = \frac{N_n(x)}{S_n(x)}$。

步骤4：计算条件期望$E[\hat{m}_h(x)|\{X_i\}]$

给定$\{X_i\}$，Y_i与X_i相关，但Y_i之间条件独立。

因此，

$$E[N_n(x)|\{X_i\}] = \sum_{i=1}^{n} K_h(x - X_i)E[Y_i|X_i] = \sum_{i=1}^{n} K_h(x - X_i)m(X_i)$$

因此，给定$\{X_i\}$，$\hat{m}_h(x)$的条件期望为：

$$E[\hat{m}_h(x)|\{X_i\}] = \frac{\sum_{i=1}^{n} K_h(x - X_i)m(X_i)}{\sum_{i=1}^{n} K_h(x - X_i)}$$

步骤5：计算无条件期望$E[\hat{m}_h(x)]$

现在，我们需要对$\{X_i\}$取期望，即：

$$E[\hat{m}_h(x)] = E_{\{X_i\}}\left[E[\hat{m}_h(x)|\{X_i\}]\right] = E_{\{X_i\}}\left[\frac{\sum_{i=1}^{n} K_h(x - X_i)m(X_i)}{\sum_{i=1}^{n} K_h(x - X_i)}\right]$$

这仍然是一个复杂的期望，因为$\{X_i\}$是随机变量。

步骤6：利用大数定律和逼近

当样本容量n趋于无穷大时，根据大数定律，我们可以近似地认为：

- 分子和分母都可以用其期望来近似；
- 或者说，经验分布可以近似真实分布。

为了简化分析，我们考虑n很大，且核密度估计的一般性质。

步骤7：引入连续性假设和核密度估计的性质

当$n \to \infty$且$h \to 0$，同时满足$nh \to \infty$，核密度估计将收敛于真实的回归函数值$m(x)$。

然而，这涉及到了渐近性质，并不能直接用于证明$E[\hat{m}_h(x)] = E[Y]$。

步骤8：考虑简单的特殊情况

为了更清楚地理解问题，我们可以考虑一个特殊情况：假设X和Y独立。

在这种情况下，$m(X) = E[Y]$是常数。

因此，

$$E[\hat{m}_h(x) | \{X_i\}] = E[Y]$$

因为$m(X_i) = E[Y]$，所以：

$$E[\hat{m}_h(x) | \{X_i\}] = \frac{\sum_{i=1}^n K_h(x - X_i) E[Y]}{\sum_{i=1}^n K_h(x - X_i)} = E[Y]$$

因此，

$$E[\hat{m}_h(x)] = E_{\{X_i\}}[E[Y]] = E[Y]$$

所以在X和Y独立的情况下，$E[\hat{m}_h(x)] = E[Y]$。

步骤9：回到一般情况

在一般情况下，$m(X_i)$不是常数，无法简单地将其移出分子。

因此，除非在特殊情况下，$E[\hat{m}_h(x)]$一般不等于$E[Y]$。

步骤10：总结

通过上述分析，我们发现：

- **在X和Y独立的特殊情况下**，$E[\hat{m}_h(x)] = E[Y]$成立。
- **在一般情况下**，由于$m(X_i)$依赖于X_i，$E[\hat{m}_h(x)]$通常不等于$E[Y]$。

结论：

一般而言，$E[\hat{m}_h(x)] \neq E[Y]$。

原因：

- 在一般情况下，$E[\hat{m}_h(x)]$是$m(X_i)$的加权平均，权重取决于核函数K_h和样本点X_i。
- 除非$m(X_i)$为常数（即Y与X独立），否则无法将$m(X_i)$移出分子的求和，从而无法简化为$E[Y]$。

因此，除非在特殊情况下，$E[\hat{m}_h(x)]$不等于$E[Y]$。

抱歉之前的疏忽，感谢您的指正！