假设事件 A 发生的概率为ε，不发生的概率为$1-\varepsilon$，重复 n 次事件 A，那么 n 次事件中事件 A 发生的次数为$G(n)$

$$P(G(n)=k)=\binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k}$$

根据霍夫丁不等式，n 次事件中 A 发生的次数小于 k 的概率的上届为：

$$P(G(n)\leq k)\leq e^{-2n(p-\frac{k}{n})^2}$$

将以上模型应用到 Bagging 中来，

对于单个基学习器而言，假设其在单个样本上预测错误(事件 A)的概率为

$$P(h_i(x)\neq f(x))=\varepsilon$$

假设有 n 个基学习器，且它们预测正确与否互不影响(独立)，那么集成后的学习器(投票机制)预测错误的概率为：

$$\begin{aligned} P\big(H(x)\neq f(x)\big) &= \sum_{k=1}^{\left\lfloor\frac{n}{2}\right\rfloor}\binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k} \\ &\leq P\left(G(n)\leq\frac{n}{2}\right) \\ &\leq e^{-2n(\varepsilon-\frac{1}{2})^2} \\ &= e^{-\frac{n}{2}(2\varepsilon-1)^2} \end{aligned}$$

可以看到在基学习器独立的条件下，随着基学习器数目 n 的增大，错误发生的概率呈指数趋势下降。