We know that $(M+1)\cdot[G(x)\neq y]\leq\sum_{t=1}^{2M+1}[g_t(x)\neq y]$ holds for any (x,y)

because when $G(x)=y$, $(M+1)\cdot 0\leq\sum_{t=1}^{2M+1}[g_t(x)\neq y]$

when $G(x)\neq y$, $M+1\leq\sum_{t=1}^{2M+1}[g_t(x)\neq y]$

Thus $E_{x\sim P}((M+1)\cdot[G(x)\neq y])\leq E_{x\sim P}(\sum_{t=1}^{2M+1}[g_t(x)\neq y])$

$$(M+1)E_{out}(G)\leq\sum_{t=1}^{2M+1}e_t$$

$$E_{out}(G)\leq\frac{\sum_{t=1}^{2M+1}e_t}{M+1}$$

Although this upper bound is valid, it is not tight.

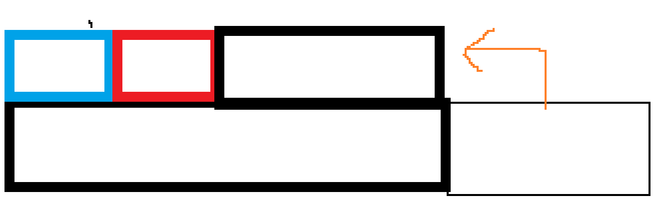
For example when $M=1$ and $e_1=e_2=e_3=1$, $\frac{\sum_{t=1}^{2M+1}e_t}{M+1}=1.5>1$, but the tightest upper bound is 1.

When $M=1$ and $e_1=e_2=0.1$, $e_3=0.6$, $\frac{\sum_{t=1}^{2M+1}e_t}{M+1}=0.4$, but the tightest upper bound is 0.2.

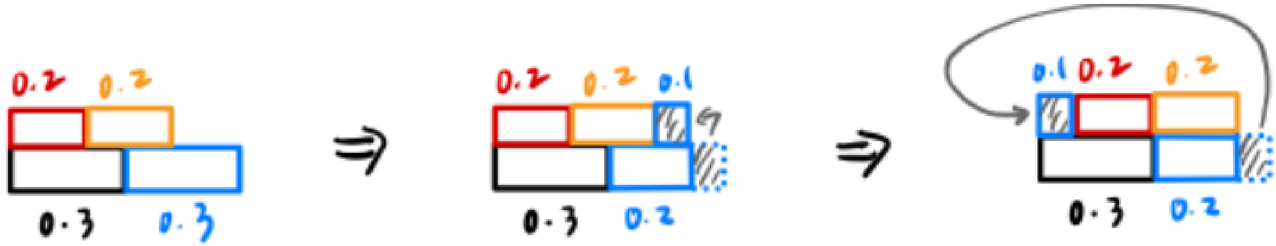
We can view this problem as a **brick stacking problem** where each e_t is a brick of length e_t and height 1. And the tightest upper bound is the maximum coverage of bricks where total height is at least $M+1$.

For example, from the previous example, we can draw:

where the maximum coverage is 0.2. The reason why $\frac{\sum e_t}{M+1}$ is not tight is because it did a cut-long-fill-short and made one classifier answer incorrectly more than once on an example:



A cut-long-fill-short is only legal if the resulting average length is longer than the brick being cut, such that it doesn't answer incorrectly more than once on an example:



With this view, the tightest upper bound should be $\min \frac{\sum_{t=1}^{i} e_t}{i-M}$

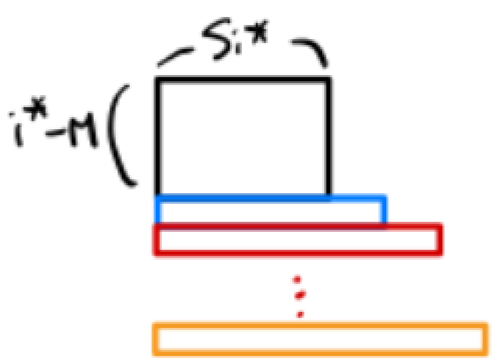
where $i = M+1, M+2, \dots, 2M+1$ and $e_1 \le e_2 \le \cdots \le e_{2M+1}$

Proof:

WLOG, we can assume that $e_1 \le e_2 \le \cdots \le e_{2M+1}$.

We can find a brick stacking method such that the maximum height of bricks is $M+1$ and the maximum coverage is $\min \frac{\sum_{t=1}^{i} e_t}{i-M}$.

Let $\frac{\sum_{t=1}^{i} e_t}{i-M} = S_i$ and $i = i^*$ when S_i has minimum, then the stacking is simply aligning the smallest i^* bricks to the left and use cut-long-fill-short to form a S_{i^*} by $i^* - M$ rectangle. And below the rectangle are $2M+1-i^*$ bricks all aligned to the left with increasing length:



This stacking is possible because:

$$\begin{cases} e_i \ge S_{i^*} & \text{for } i > i^* & \dots① \\ e_i \le S_{i^*} & \text{for } i \le i^* & \dots② \end{cases}$$

Proof for ①:

$e_i \ge S_{i^*}$ for $i > i^* \iff e_{i^*+1} \ge S_{i^*}$

$$S_{i^*+1} = \frac{\sum_{t=1}^{i^*+1} e_t}{i^*+1-M} \geq \frac{\sum_{t=1}^{i^*} e_t}{i^*-M} = S_{i^*}$$

$$\frac{e_{i^*+1}}{i^*+1-M} + S_{i^*} \cdot \frac{i^*-M}{i^*+1-M} \geq S_{i^*}$$

$$\frac{e_{i^*+1}}{i^*+1-M} \geq \frac{S_{i^*}}{i^*+1-M}$$

$$e_{i^*+1} \geq S_{i^*} \quad (Q.E.D)$$

Proof for ②:

$e_i \leq S_{i^*}$ for $i \leq i^* \iff e_{i^*} \leq S_{i^*}$

$$S_{i^*} = \frac{\sum_{t=1}^{i^*} e_t}{i^*-M} \leq \frac{\sum_{t=1}^{i^*-1} e_t}{i^*-1-M} = S_{i^*-1}$$

$$\frac{e_{i^*}}{i^*-M} + S_{i^*-1} \cdot \frac{i^*-1-M}{i^*-M} \leq S_{i^*-1}$$

$$\frac{e_{i^*}}{i^*-M} \leq \frac{S_{i^*-1}}{i^*-M}$$

$$e_{i^*} \leq S_{i^*-1}$$

Also since $S_{i^*} = (\frac{1}{i^*-M})e_{i^*} + (1-\frac{1}{i^*-M})S_{i^*-1}$ and $e_{i^*} \leq S_{i^*-1}$

$$\therefore e_{i^*} \leq S_{i^*} \quad (Q.E.D)$$

This stacking is optimal because for S_i where $i > i^*$, an illegal cut-long-fill-short occurs. For S_i where $i < i^*$, there are bricks below the rectangle that are shorter than the rectangle, which are all invalid.

Thus the tightest upper bound is $\min \frac{\sum_{t=1}^{i} e_t}{i-M}$

where $i = M+1, M+2, \ldots, 2M+1$ and $e_1 \leq e_2 \leq \cdots \leq e_{2M+1}$