- Suppose $\varepsilon_j \sim N(\mu, \sigma^2)$, $j = 1, \ldots, J$, $J \geq 2$
- Question: What is $E[\varepsilon_j]$? Ans: $E[\varepsilon_j] = \mu$ ✓

What is $E\left[\min_{1 \leq j \leq J} \varepsilon_j\right]$? $\geq \mu$ ×, $= \mu$ ×, $< \mu$ ✓

$= E[\varepsilon_{j^*}]$

$E[\varepsilon_{j^*}] < \mu = \mu_{j^*}$

$j^* = \arg\min_{1 \leq j \leq J} \varepsilon_j$

$\varepsilon_j \equiv$ CV error of j^{th} model

e.g. $\varepsilon_1 \equiv$ CV error of logistic regression

$\varepsilon_2 \equiv$ " " " random forests

$\varepsilon_3 \equiv$ " " " 5-NN

In practice $\varepsilon_j \sim N(\mu_j, \sigma_j^2)$

But still $E[\varepsilon_{j^*}] < \mu_{j^*}$