Ex. 15.4 Suppose $x_i,\ i = 1, \ldots, N$ are iid (μ, σ^2). Let $\bar{x}_1^*$ and $\bar{x}_2^*$ be two bootstrap realizations of the sample mean. Show that the sampling correlation $\text{corr}(\bar{x}_1^*, \bar{x}_2^*) = \frac{n}{2n-1} \approx 50\%$. Along the way, derive $\text{var}(\bar{x}_1^*)$ and the variance of the bagged mean $\bar{x}_{bag}$. Here $\bar{x}$ is a *linear* statistic; bagging produces no reduction in variance for linear statistics.

First, note that $E[x_i^*] = \boxed{\mu}$

because bootstrap samples have the same expected value as the original sample.

Then, $VAR[x_i^*] = E[VAR[x_i^* \mid x_1 \ldots x_N]] + VAR[E[x_i^* \mid x_1 \ldots x_n]]$

$$= E\left[\frac{N-1}{N} s^2\right] + VAR[\bar{x}]$$

$$= \frac{N-1}{N}\sigma^2 + \frac{\sigma^2}{N} = \boxed{\sigma^2}$$

So then $VAR[\bar{x}^* \mid x_1 \ldots x_n]$

$$= \frac{1}{N^2} VAR\left(\sum x_i^* \mid x_1 \ldots x_n\right)$$

$$= \frac{2N-1}{N} \frac{s^2}{N} = \boxed{\frac{(2N-1)\, s^2}{N^2}}$$

$$cov(\bar{x}_1^*, \bar{x}_2^*) = \frac{1}{n^2} \sum_{ij} cov(x_i, x_j) = \frac{s^2}{N}$$

So $$Corr(\bar{x}_1^*, \bar{x}_2^*) = \frac{cov}{VAR} = \frac{\frac{s^2}{N}}{\frac{N-1\, s^2}{N^2}} = \boxed{\frac{N}{2N-1}}$$