Ex. 15.4 Suppose x_i, $i = 1, \dots, N$ 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 $\mathrm{corr}(\bar{x}_1^*, \bar{x}_2^*) = \frac{n}{2n-1} \approx 50\%$. Along the way, derive $\mathrm{var}(\bar{x}_1^*)$ 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, $VAR[x_i^*] = E[VAR[x_i^* \mid x_1 \dots x_N]] + VAR[E[x_i^* \mid x_1 \dots x_n]]$

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

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

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

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

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

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

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

If B bootstrap replicates, then

$$\mathrm{Var}(\bar{X}_{Bag}) = \mathrm{Var}\left(\frac{1}{B}\sum_{i=1}^{B} \bar{X}_i^*\right)$$

$$= \frac{1}{B^2}\sum_{i=1}^{B} \mathrm{VAR}(\bar{X}_i^*) + \frac{1}{B^2}\,\mathrm{Cov}(\bar{X}_j^*, \bar{X}_k^*)$$

$$= \frac{1}{B}\cdot\frac{(2n-1)\,\sigma^2}{n^2} + \frac{B-1}{B}\cdot\frac{\sigma^2}{n}$$

$$= \boxed{\frac{\sigma^2\left[(2n-1)+(B-1)n\right]}{Bn^2}}$$

3