Problem 1.

Writing out $\hat{\theta}$, we have

$$\begin{aligned}\hat{\theta} &= \hat{F}_n(b) - \hat{F}_n(a) \\ &= \frac{1}{n}\sum_{i=1}^{n} 1\{X_i \le b\} - \frac{1}{n}\sum_{i=1}^{n} 1\{X_i \le a\} \\ &= \frac{1}{n}\sum_{i=1}^{n} \left(1\{X_i \le b\} - 1\{X_i \le a\}\right).\end{aligned}$$

What are the potential possibilities for the indicators? We have

$$\begin{cases} X_i \le b \text{ and } X_i \le a & \implies 1\{X_i \le b\} - 1\{X_i \le a\} = 1 - 1 = 0 \\ X_i \le b \text{ but } X_i > a & \implies 1\{X_i \le b\} - 1\{X_i \le a\} = 1 - 0 = 1 \\ X_i > b \text{ but } X_i \le a & \text{this is impossible, since } a < b \\ X_i > b \text{ and } X_i > a & \implies 1\{X_i \le b\} - 1\{X_i \le a\} = 0 - 0 = 0 \end{cases}.$$

We can see then that the only time the indicator equals 1 is in the second case, when $X_i \le b$ but $X_i > a$, so

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} 1\{a < X_i \le b\}.$$

Let $Y_i = 1\{a < X_i \le b\}$. Then $Y_1, Y_2, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Bern}(p)$, where p is the probability that the indicator equals 1. That is,

$$p = \mathbb{P}(Y_i = 1) = P(a < X_i \le b) = F(b) - F(a) = \theta,$$

by a property of the CDF. Altogether then, we can write the estimator as

$$\hat{\theta} = \hat{F}_n(b) - \hat{F}_n(a) = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}_n = \hat{p},$$

where $\hat{p}$ is the sample proportion of the Bernoulli RVs. Recall that for iid Bernoulli RVs, $\text{Var}(\hat{p}) = p(1-p)/n$. Thus,

$$\text{Var}(\hat{\theta}) = \frac{p(1-p)}{n} = \frac{\theta(1-\theta)}{n} = \frac{(F(b) - F(a))\,(1 - F(b) + F(a))}{n}. \tag{1}$$

Plugging in the estimator for $F(b) - F(a)$ and taking the square root, the estimated standard error is

$$\hat{se}(\hat{\theta}) = \sqrt{\frac{\left(\hat{F}_n(b) - \hat{F}_n(a)\right)\left(1 - \hat{F}_n(b) + \hat{F}_n(a)\right)}{n}}.$$

Using the asymptotic normality of the sample proportion since $\hat{\theta}_n = \hat{p}$, a $(1-\alpha)100\%$ approximate confidence interval for θ is

$$\hat{F}_n(b) - \hat{F}_n(a) \pm z_{\alpha/2}\sqrt{\frac{\left(\hat{F}_n(b) - \hat{F}_n(a)\right)\left(1 - \hat{F}_n(b) + \hat{F}_n(a)\right)}{n}}.$$

Problem 2.

For the conditional expectation, we have

$$\begin{aligned}\mathbb{E}(\bar{X}_n^*|X_1,\ldots,X_n) &= \mathbb{E}\left(\frac{1}{n}\sum_{j=1}^{n} X_j^* \middle| X_1,\ldots,X_n\right)\\ &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j^*|X_1,\ldots,X_n) \text{ (by linearity of expectation).}\end{aligned}$$

X_j^* is one of the RVs in the bootstrap sample, and it has conditional expectation

$$\mathbb{E}(X_j^*|X_1,\ldots,X_n) = \sum_{t\in\mathcal{T}} t\ \mathbb{P}(X_j^* = t|X_1,\ldots,X_n),$$

where $\mathcal{T}$ is the set of possible values that X_j^* can take on. But since the data are unique, $\mathcal{T}$ is just $\{X_1,\ldots,X_n\}$. Meanwhile, the conditional probability that X_j^* takes on value t is just $1/n$, since a bootstrapped sample draws n observations with replacement and with equal probability. So we have

$$\mathbb{E}(X_j^*|X_1,\ldots,X_n) = \frac{1}{n}\sum_{k=1}^{n} X_k = \bar{X}_n,$$

and

$$\mathbb{E}(\bar{X}_n^*|X_1,\ldots,X_n) = \frac{1}{n}\sum_{j=1}^{n}\bar{X}_n = \frac{1}{n}\cdot n\bar{X}_n = \bar{X}_n.$$

Similarly for the variance, we have

$$\text{Var}(\bar{X}_n^*|X_1,\ldots,X_n) = \frac{1}{n^2}\sum_{j=1}^{n}\text{Var}(X_j^*|X_1,\ldots,X_n),$$

which follows by a property of variance, since the bootstrap members are iid.

$$\text{Var}(X_j^*|X_1,\ldots,X_n) = \sum_{t\in\tilde{\mathcal{T}}}(t-\mathbb{E}(X_j^*|X_1,\ldots,X_n))^2\mathbb{P}(X_j^* = t|X_1,X_2,\ldots,X_n),$$

where likewise, the set of possible values is $\tilde{\mathcal{T}} = \{X_1,\ldots,X_n\}$ and the probability is $1/n$. Plugging in the conditional expectation too, we have

$$\text{Var}(X_j^*|X_1,\ldots,X_n) = \frac{1}{n}\sum_{k=1}^{n}(X_i-\bar{X})^2,$$

so

$$\begin{aligned}\mathrm{Var}(\bar{X}_n^*|X_1,\ldots,X_n) &= \frac{1}{n^2}\sum_{j=1}^{n}\left(\frac{1}{n}\sum_{k=1}^{n}(X_i-\bar{X})^2\right)\\ &= \frac{1}{n^2}\cdot n\left(\frac{1}{n}\sum_{k=1}^{n}(X_i-\bar{X})^2\right)\\ &= \frac{1}{n^2}\sum_{k=1}^{n}(X_i-\bar{X})^2.\end{aligned}$$

Let $\mu = \mathbb{E}(X_i)$ and $\sigma^2 = \mathrm{Var}(X_i)$.

To get the unconditional mean of the bootstrap sample, let us apply the Law of Total Expectation:

$$\begin{aligned}\mathbb{E}(\bar{X}_n^*) &= \mathbb{E}\left(\mathbb{E}(\bar{X}_n^*|X_1,\ldots,X_n)\right)\\ &= \mathbb{E}(\bar{X}_n) = \mathbb{E}(X_i) = \mu,\end{aligned}$$

Similarly applying the Law of Total Variance, we have

$$\begin{aligned}\mathrm{Var}(\bar{X}_n^*) &= \mathbb{E}\left(\mathrm{Var}(\bar{X}_n^*|X_1,\ldots,X_n)\right) + \mathrm{Var}\left(\mathbb{E}(\bar{X}_n^*|X_1,\ldots,X_n)\right)\\ &= \mathbb{E}\left(\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\bar{X})^2\right) + \mathrm{Var}(\bar{X}).\end{aligned}$$

The first term can be computed by noting that

$$\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{\hat{\sigma}^2}{n} = \frac{n-1}{n^2}s^2,$$

where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$, the unbiased estimator of σ^2. So

$$\mathbb{E}\left(\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\bar{X})^2\right) = \frac{(n-1)}{n^2}\mathbb{E}(s^2) = \frac{n-1}{n^2}\sigma^2.$$

And clearly $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$. Finally then,

$$\begin{aligned}\mathrm{Var}(\bar{X}_n^*) &= \frac{n-1}{n^2}\sigma^2 + \frac{\sigma^2}{n}\\ &= \frac{(n-1)\sigma^2 + n\sigma^2}{n^2}\\ &= \frac{2n-1}{n^2}\sigma^2\end{aligned}$$