Chapter 8 Solutions

Andrew Wu
Wasserman: All of Statistics

April 1, 2025

Problem 8.4. Let $X_1, \ldots, X_n$ be distinct observations (no ties.) Show that there are $\binom{2n-1}{n}$ distinct bootstrap samples.

Solution. We want to compute the number of ways to select n of $X_1, \ldots, X_n$ with replacement.

Imagine instead that we have n stars, and $n-1$ bars, and that we order them in a row. We claim any distinct arrangement is equivalent to a distinct bootstrap sample.

Given an arrangement of stars and bars, we will say that the number of times X_i is picked is equal to the number of stars between bar $i-1$ and bar i (with the caveat that X_1 is picked k times, where k is the number of stars before the first bar, and that X_n is picked j times, where j is the number of stars after the last bar.) Moreover, given a bootstrap sample, we can replicate the procedure in reverse.

Therefore there is a correspondence between distinct arrangements and distinct bootstrap samples, and so there are $\binom{2n-1}{n}$ distinct bootstrap samples. □

Problem 8.5. Let $X_1, \ldots, X_n$ be distinct observations (no ties.) Let $X_1^*, \ldots, X_n^*$ denote a bootstrap sample, and let $\overline{X}_n^* = \frac{1}{n}\sum_{i=1}^n X_i^*$. Find the following:

- $\mathbb{E}(\overline{X}_n^* | X_1, \ldots, X_n)$,
- $\mathbb{V}(\overline{X}_n^* | X_1, \ldots, X_n)$,
- $\mathbb{E}(\overline{X}_n^*)$,
- $\mathbb{V}(\overline{X}_n^*)$.

Solution. We have

$$\begin{aligned}
\mathbb{E}(\overline{X}_n^* | X_1, \ldots, X_n) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i^* \middle| X_1, \ldots, X_n\right) \\
&= \frac{1}{n} \cdot n \cdot \mathbb{E}(X_i^* | X_1, \ldots, X_n) \\
&= \overline{X}_n.
\end{aligned}$$

And for the variance,

$$\begin{aligned}
\mathbb{V}(\overline{X}_n^* | X_1, \ldots, X_n) &= \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^n X_i^* \middle| X_1, \ldots, X_n\right) \\
&= \frac{1}{n^2} \cdot n \cdot \mathbb{V}(X_i^* | X_1, \ldots, X_n) \\
&= \frac{1}{n}[(\mathbb{E}((X_i^*)^2) | X_1, \ldots, X_n) - (\mathbb{E}(X_i^*) | X_1, \ldots, X_n)^2] \\
&= \frac{1}{n}\left[\frac{1}{n}(X_1^2 + \ldots + X_n^2) - \overline{X}_n^2\right]
\end{aligned}$$

Next, we have

$$\begin{aligned}
\mathbb{E}(\overline{X}_n^*) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i^*\right) \\
&= \frac{1}{n}\cdot n\cdot \mathbb{E}(X_i^*) \\
&= \mathbb{E}(X_1).
\end{aligned}$$

And finally,

$$\begin{aligned}
\mathbb{V}(\overline{X}_n^*) &= \mathbb{E}(\mathbb{V}(\overline{X}_n^*|X_1,\dots,X_n)) + \mathbb{V}(\mathbb{E}(\overline{X}_n^*|X_1,\dots,X_n)) \\
&= \mathbb{E}\left(\frac{1}{n}\left(\frac{1}{n}\left(X_1^2+\ldots+X_n^2\right) - \overline{X}_n^2\right)\right) + \mathbb{V}(\overline{X}_n) \\
&= \frac{1}{n}\left[\mathbb{E}\left(\frac{1}{n}(X_1^2+\ldots+X_n^2)\right) - \mathbb{E}\left(\overline{X}_n^2\right)\right] + \frac{1}{n^2}\cdot n\cdot \mathbb{V}(X_1) \\
&= \frac{1}{n}\left[\frac{1}{n}\cdot n\cdot \mathbb{E}(X_1^2) - \mathbb{V}(\overline{X}_n) - \mathbb{E}(\overline{X}_n)^2\right] + \frac{1}{n}\mathbb{V}(X_1) \\
&= \frac{1}{n}\left(\mathbb{E}(X_1^2) - \frac{1}{n}\mathbb{V}(X_1) - \mathbb{E}(X_1)^2\right) + \frac{1}{n}\mathbb{V}(X_1) \\
&= \frac{1}{n}\cdot\frac{n-1}{n}\cdot \mathbb{V}(X_1) + \frac{1}{n}\mathbb{V}(X_1) \\
&= \frac{2n-1}{n^2}\mathbb{V}(X_1).
\end{aligned}$$

□

Problem 8.7. Let $X_1,\dots,X_n \sim \text{Uniform}(0,\theta)$. Let $\widehat{\theta} = \max\{X_1,\dots,X_n\}$. With $\theta = 1$ and $n = 50$, find the distribution of $\widehat{\theta}$. Show that if $\widehat{\theta}^*$ is a bootstrapped estimate for $\widehat{\theta}$, then $\mathbb{P}(\widehat{\theta}^* = \widehat{\theta}) \approx 0.632$.

Solution. We have, for $0 \le x \le \theta$,

$$\begin{aligned}
\mathbb{P}(\widehat{\theta} \le x) &= \mathbb{P}(X_1,\dots,X_n \le x) \\
&= \mathbb{P}(X_1 \le x)^n \\
&= \left(\frac{x}{\theta}\right)^n.
\end{aligned}$$

Thus $F(x) = \left(\frac{x}{\theta}\right)^n$ for $0 \le x \le \theta$, so $f(x) = F'(x) = \frac{1}{\theta^n}nx^{n-1}$. Taking $\theta = 1$ and $n = 50$, we obtain $f(x) = 50x^{49}$.

Now, if $\widehat{\theta}^*$ is a bootstrapped estimate for θ, we find $\widehat{\theta}^*$ by picking n times from $\{X_1,\dots,X_n\}$ with replacement, and then taking the maximum of those picks. Without loss of generality, assume $X_n = \max\{X_1,\dots,X_n\}$. Then, $\mathbb{P}(\widehat{\theta}^* \ne \widehat{\theta})$ is the probability that none of the picks are X_n, so thus $\mathbb{P}(\widehat{\theta}^* \ne \widehat{\theta}) = (\frac{n-1}{n})^n$. It follows that

$$\begin{aligned}
\mathbb{P}(\widehat{\theta}^* = \widehat{\theta}) &= 1 - \left(\frac{n-1}{n}\right)^n \\
&= 1 - \left(1 - \frac{1}{n}\right)^n \\
&\approx 1 - \frac{1}{e} \approx 0.632.
\end{aligned}$$

□

Problem 8.8. Let $T_n = \overline{X}_n^2$, $\mu = \mathbb{E}(X_1)$, $\alpha_k = \int |x-\mu|^k dF(x)$, and $\widehat{\alpha}_k = n^{-1}\sum_{i=1}^n |X_i - \overline{X}_n|^k$. Show that

$$v_{\text{boot}} = \frac{4\overline{X}_n^2\widehat{\alpha}_2}{n} + \frac{4\overline{X}_n\widehat{\alpha}_3}{n^2} + \frac{\widehat{\alpha}_4}{n^3}.$$

Solution. See this StackExchange post; I was unable to solve this problem myself.

□