6. A confidence interval for uniform distributions

(a)

2.0/2 points (graded)

Let $X_1, \ldots, X_n$ be i.i.d. uniform random variables in $[0, \theta]$, for some $\theta > 0$. Denote by

$$M_n = \max_{i=1,\ldots,n} X_i.$$

Compute the following probabilities:

$\mathbf{P}(M_n \geq \theta) =$ 0 ✔ **Answer:** 0

0

For all $0 \leq t \leq \theta$:

$\mathbf{P}(M_n \leq \theta - t) =$ ((theta - t)/theta)^n ✔ **Answer:** (1 - t/theta)^n

(*Food for thought:* What can you conclude?)

STANDARD NOTATION

Solution:

First, $M_n \leq \theta$ almost surely, because all $X_i \leq \theta$ almost surely, so

$$\mathbf{P}(M_n \geq \theta) = 0.$$

Second, let $0 \leq t \leq \theta$. Because having an upper bound on the maximum of n variables is the same as having an upper bound on all of the variables, and the X_i are independent, we can write

$$\begin{aligned}
\mathbf{P}(M_n \leq \theta - t) = \;& \mathbf{P}(X_i \leq \theta - t \text{ for all } i = 1, \ldots, n) && \\
= \;& \prod_{i=1}^{n} \mathbf{P}(X_i \leq \theta - t) && \text{(by independence)} \\
= \;& (\mathbf{P}(X_1 \leq \theta - t))^n && \text{(all } X_i \text{ have the same distribution)} \\
= \;& \left(\frac{\theta - t}{\theta}\right)^n && \text{(cdf of Uniform distribution)} \\
= \;& \left(1 - \frac{t}{\theta}\right)^n \xrightarrow[n\to\infty]{} 0 &&
\end{aligned}$$

Hence,

$$M_n \xrightarrow{\mathbf{P}} \theta.$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

(b)

2.0/2 points (graded)

Compute the cumulative distribution function $F_n(t)$ of $n(1 - M_n/\theta)$ for fixed $t \in [0, n]$ and any positive integer n.

$F_n(t) =$ 1-(1-t/n)^n **Answer:** 1-(1-t/n)^n

Compute the following limit.

$\lim_{n\to\infty} F_n(t) =$ 1-exp(-t) **Answer:** 1 - exp(-t)

(*Food for thought:* Again, What can you conclude?)

STANDARD NOTATION

Solution:

Let $t > 0$ and first observe that we can rewrite

$$n\left(1 - \frac{M_n}{\theta}\right) \le t \iff M_n \ge \theta - \theta\frac{t}{n}.$$

For n large enough, $t/n \le 1$. Together with the fact that the cdf of M_n does not have atoms, we can compute:

$$\begin{aligned}
\mathbf{P}\left(n\left(1 - \frac{M_n}{\theta}\right) \le t\right) &= \mathbf{P}\left(M_n \ge \theta - \theta\frac{t}{n}\right) \\
&= 1 - \mathbf{P}\left(M_n \le \theta - \theta\frac{t}{n}\right) \\
&= 1 - \left(1 - \frac{t}{n}\right)^n \qquad \text{(by part (a))} \\
&\xrightarrow[n\to\infty]{} 1 - \exp(-t).
\end{aligned}$$

To obtain the limit, we used the limit formula for the exponential,

$$\left(1 + \frac{a}{n}\right)^n \xrightarrow[n\to\infty]{} \exp(a), \quad \text{for } a \in \mathbb{R}.$$

Therefore,

$$n(1 - M_n/\theta) \xrightarrow[n\to\infty]{(D)} \mathrm{Exp}(1),$$

that is, it converges to an Exponential random variable with parameter 1.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

(c)

1/2 points (graded)

Next, we will use the previous question to find an interval $\mathcal{I}$ of the form $\mathcal{I} = [M_n, M_n + c]$, that does not depend on θ and such that

$$\mathbf{P}\left[\mathcal{I} \ni \theta\right] \rightarrow .95, \text{ as } n \rightarrow \infty.$$

The strategy now is to use a plug-in estimator for θ to replace it in the expression for c. Parts (a) and (b) suggest that we use c of the form $\left(\frac{t}{n}\right) M_n$, where t ought to equal a certain value in order for $\mathbf{P}\left[\mathcal{I} \ni \theta\right] \rightarrow .95$. What is the appropriate numerical value of t?

$t =$ ln(0.05) ✖ **Answer:** ln(20)

Why can we use a plugin-estimator for the asymptotic confidence interval?

- ○ By the Delta Method, the asymptotic variance scales with the square of the first derivative of the plugin function.
- ◉ By Slutsky's Theorem, we can combine convergence in distribution of Y_n and in probability of Z_n if Z_n converges to a constant. ✔
- ○ By the Central Limit Theorem, the plugin variable will again be normally distributed.

STANDARD NOTATION

忘了前面还有一个负号

Solution:

Here is a presentation of the argument, in full. In summary, $t = \log(20)$ due to the fact that we want $0.95 = 1 - \exp(-t)$. We arrive at this conclusion via Slutsky's theorem.

By part (a), we know that $\theta \geq M_n$ almost surely. Moreover, for any $t > 0$, by part (b), we have that

$$\mathbf{P}\left(\theta \geq M_n + \theta\frac{t}{n}\right) \xrightarrow[n\to\infty]{} \exp(-t).$$

Moreover, by part (a), we know that

$$M_n \xrightarrow{\mathbf{P}} \theta,$$

which is a constant. By Slutsky's Theorem, we can substitute M_n for θ above to obtain

$$\mathbf{P}\left(\theta \geq M_n + M_n\frac{t}{n}\right) \xrightarrow[n\to\infty]{} \exp(-t).$$

Pick

$$t = \log(20)$$

and set

$$\mathcal{I} = \left[M_n, M_n + M_n\frac{\log(20)}{n}\right]$$

With this, we obtain

$$
\begin{aligned}
\mathbf{P}(\mathcal{I} \ni \theta) = \; & 1 - \underbrace{\mathbf{P}(\theta \leq M_n)}_{=0} - \mathbf{P}\left(\theta \geq M_n + M_n \frac{\log(20)}{n}\right) \\
\rightarrow \; & 1 - \exp(-\log(20)) = 0.95.
\end{aligned}
$$

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(d)

0/1 point (graded)

Compute the bias of M_n as an estimator of θ.

$\mathbb{E}[M_n] - \theta =$ -1/(n+1) ✖ **Answer:** -theta/(n+1)

$-\frac{1}{n+1}$

STANDARD NOTATION

Solution:

By part (a), we know that for $r \in [0, \theta]$,

$$
\mathbf{P}(M_n \leq r) = \left(1 - \frac{\theta - r}{\theta}\right)^n = \left(\frac{r}{\theta}\right)^n,
$$

and that the support of M_n is $[0, \theta]$. Hence, the density f_n of M_n is

$$
f_n(r) = \begin{cases} 0, & r < 0 \text{ or } r > \theta \\ \frac{1}{\theta} n \left(\frac{r}{\theta}\right)^{n-1}, & 0 \leq r \leq \theta \end{cases}
$$

Therefore, we can compute its expectation,

$$
\begin{aligned}
\mathbb{E}[M_n] &= \int_0^\theta \frac{nr}{\theta} \left(\frac{r}{\theta}\right)^{n-1} dr \\
&= \frac{n}{(n+1)\theta^n} r^{n+1} \Big|_0^\theta = \frac{n}{n+1}\theta.
\end{aligned}
$$

That means that the bias of M_n is

$$
\mathbb{E}[M_n] - \theta = -\frac{1}{n+1}\theta.
$$

If we wanted, we could therefore obtain an unbiased estimator $\tilde{M}_n$ by setting

$$
\tilde{M}_n = \frac{n+1}{n} M_n.
$$

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem