10. Confidence Interval for an Exponential Statistical Model

Confidence Interval for an Exponential Statistical Model

So you have that lambda is between lambda 1

plus lambda hat plus lambda q alpha over 2 divided by root n,

lambda hat minus lambda q alpha over 2 divided by root n.

I put on my plug-in hat, and there you go.

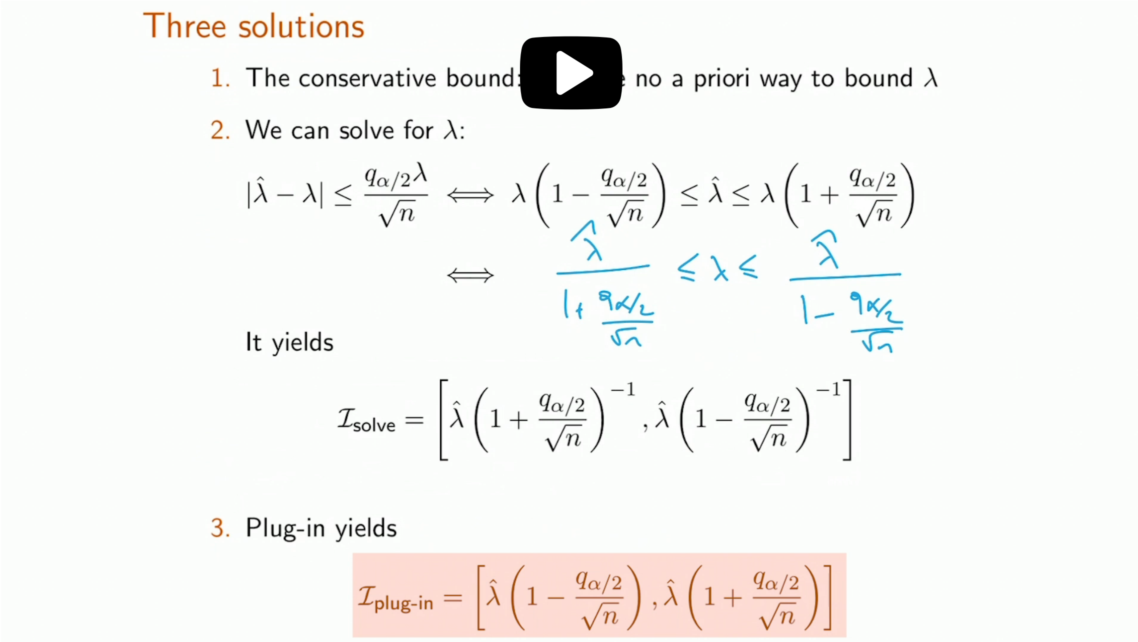
I have a new confidence interval which is my plug-in confidence

interval.

I factor out my lambda hat, and I

get something which is centered about lambda hat--

much easier.



5:30 / 5:30 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Confidence interval Concept Check

1/1 point (graded)

As in the previous section, let $X_1, \ldots, X_n \overset{iid}{\sim} \exp(\lambda)$. Let

$$\widehat{\lambda}_n := \frac{n}{\sum_{i=1}^n X_i}$$

denote an estimator for λ. We know by now that $\widehat{\lambda}_n$ is a **consistent** and **asymptotically normal** estimator for λ.

Recall $q_{\alpha/2}$ denote the $1-\alpha/2$ quantile of a standard Gaussian. By the Delta method:

$$\lambda \in \left[\widehat{\lambda}_n - \frac{q_{\alpha/2}\lambda}{\sqrt{n}}, \widehat{\lambda}_n + \frac{q_{\alpha/2}\lambda}{\sqrt{n}}\right] =: \mathcal{I}$$

with probability $1-\alpha$. However, $\mathcal{I}$ is still **not** a confidence interval for λ.

Why is this the case?

- $\mathcal{I}$ is actually a confidence interval for $1/\lambda$, not λ.
- The endpoints of $\mathcal{I}$ depend on the true parameter.
- A confidence interval is supposed to be random, but $\mathcal{I}$, as constructed, is not.
- As written, the left endpoint of $\mathcal{I}$ may be larger than the right endpoint of $\mathcal{I}$, in which case $\mathcal{I}$ would not even be a valid interval.

✔

Solution:

The **second choice** is correct. The expression for the left and right endpoint of $\mathcal{I}$ both depend on the true parameter λ. By definition, a confidence interval must be computed only using the data and other known quantities, but not the true parameter, which is unknown. Therefor $\mathcal{I}$ is not a valid confidence interval.

Now we examine the incorrect choices.

- The first choice '$\mathcal{I}$ is actually a confidence interval for $1/\lambda$, not λ' is incorrect because, as already discussed, $\mathcal{I}$ cannot be a confidence in the first place because its endpoints depend on the true parameter.

- The third choice 'A confidence interval is supposed to be random, but $\mathcal{I}$, as constructed, is not' is also incorrect. The randomness for $\mathcal{I}$ comes from $\widehat{\lambda}_n$, which is random because it depends on the sample. Recall that $\widehat{\lambda}_n$ is the reciprocal of the sample mean.

- The fourth choice 'As written, the left endpoint of $\mathcal{I}$ may be larger than the right endpoint of $\mathcal{I}$, in which case $\mathcal{I}$ would not be a valid interval' is also incorrect. Since $q_{\alpha/2}$, λ, and $\sqrt{n}$ are all positive numbers, it follows that

$$\widehat{\lambda}_n - \frac{q_{\alpha/2}\lambda}{\sqrt{n}} < \widehat{\lambda}_n + \frac{q_{\alpha/2}\lambda}{\sqrt{n}}.$$

  Hence, $\mathcal{I}$ is always a valid interval, just not a valid **confidence** interval.

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Conservative confidence interval for a exponential model

1/1 point (graded)

This problem illustrates the failure of the 'conservative method' for constructing confidence intervals for an exponential statistical model.

As above, let $X_1, \ldots, X_n \stackrel{iid}{\sim} \exp(\lambda)$. Let

$$\widehat{\lambda}_n := \frac{n}{\sum_{i=1}^n X_i}$$

denote an estimator for λ.

Previously, we used the Delta method to show that for n sufficiently large

$$\lambda \in \left[\widehat{\lambda}_n - \frac{q_{\alpha/2}\lambda}{\sqrt{n}}, \widehat{\lambda}_n + \frac{q_{\alpha/2}\lambda}{\sqrt{n}}\right] =: \mathcal{I}$$

where $q_{\alpha/2}$ is the $1-\alpha/2$ quantile of a standard Gaussian.

Given an interval of the form $\mathcal{I}$, we can use the “conservative method" to find a confidence interval $\mathcal{I}_{cons}$ for λ defined by

$$\mathcal{I}_{cons} := \left[\widehat{\lambda}_n - \max_{\lambda\in(0,\infty)} \frac{q_{\alpha/2}\lambda}{\sqrt{n}}, \widehat{\lambda}_n + \max_{\lambda\in(0,\infty)} \frac{q_{\alpha/2}\lambda}{\sqrt{n}}\right].$$

Which of the following is $\mathcal{I}_{cons}$?

- (x) $(-\infty, \infty)$
- ( ) the empty interval
- ( ) the point $\widehat{\lambda}_n$
- ( ) Cannot be determined, since the exact form of $\mathcal{I}_{cons}$ will depend on the particular sample.

✔

Solution:

Observe that

$$\max_{\lambda\in(0,\infty)} \frac{q_{\alpha/2}\lambda}{\sqrt{n}} = \infty$$

. In this case, we take the max over the interval $(0, \infty)$, because a priori, λ can be any number in this interval. Therefore,

$$\mathcal{I}_{cons} = (-\infty, \infty) \, .$$

Remark: Although $(-\infty, \infty)$ is technically still a confidence interval (it even has level 100%!), it is not useful for statistical purposes because such a confidence interval gives no information about the location of the true parameter.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Explicit Confidence Intervals for an Exponential Statistical Model

6/6 points (graded)

Suppose you observe a sample data set consisting of $n = 64$ inter-arrival times $X_1, \ldots, X_{64}$ for the subway, measured in minutes. As before, we assume the statistical model that $X_1, \ldots, X_{64} \stackrel{iid}{\sim} \exp(\lambda)$ for some unknown parameter $\lambda > 0$. In this data set, you observe that the sample mean is $\frac{1}{64} \sum_{i=1}^{64} X_i = 7.8$.

Additional Instructions: For best results, please adhere to the following guidelines and reminders:

1. *For the upcoming calculations, please truncate $q_{\alpha/2}$ at 2 decimal places, instead of a more exact value. For example, if $q_{\alpha/2} = 3.84941$, use 3.84 instead of 3.85 or 3.849.*
2. *Input answers truncated at 4 decimal places. For example, if your calculations yield 11.327458, use 11.3274 instead of 11.3275 or 11.32745.*
3. *You will be computing CIs at **asymptotic level** 90%.*

Using the 'solve method' (**refer to the slide 'Three solutions'**), construct a confidence interval $\mathcal{I}_{solve}$ with asymptotic level 90% for the unknown parameter λ.

[0.1062 ✔ **Answer:** 0.1063408 , 0.1615 ✔ **Answer:** 0.1613875]

Using the 'plug-in method' $\mathcal{I}_{plug-in}$ (**refer to the slide 'Three solutions'**), construct a confidence interval with asymptotic level 90% for the unknown parameter λ.

[0.1017 ✔ **Answer:** 0.1018453 , 0.1546 ✔ **Answer:** 0.154565]

Which interval is narrower?

- ( ) $\mathcal{I}_{solve}$
- (•) $\mathcal{I}_{plug-in}$

✔

Which of these confidence intervals is centered about the sample estimate, $\hat{\lambda}_n$?

- ( ) $\mathcal{I}_{solve}$
- (•) $\mathcal{I}_{plug-in}$
- ( ) Both
- ( ) Neither

✔

Solution:

The formula for $\mathcal{I}_{solve}$ at asymptotic level $1-\alpha$ is given by

$$\mathcal{I}_{solve} = \left[\widehat{\lambda_n}\left(1+\frac{q_{\alpha/2}}{\sqrt{n}}\right)^{-1}, \widehat{\lambda_n}\left(1-\frac{q_{\alpha/2}}{\sqrt{n}}\right)^{-1}\right].$$

We need to construct a confidence interval of (asymptotic) level 90%, so this implies that $\alpha = 0.1$ and thus $q_{\alpha/2} = q_{0.05} \approx 1.64$ (consulting a table for the standard Gaussian). Hence, for this data set,

$$\begin{aligned}\mathcal{I}_{solve} &= \left[\frac{1}{7.8}\left(1+\frac{1.64}{\sqrt{64}}\right)^{-1}, \frac{1}{7.8}\left(1-\frac{1.64}{\sqrt{64}}\right)^{-1}\right] \\ &\approx [0.1064, 0.1613].\end{aligned}$$

Next we compute $\mathcal{I}_{plug-in}$. The formula is given by

$$\mathcal{I}_{plug-in} = \left[\widehat{\lambda_n}\left(1-\frac{q_{\alpha/2}}{\sqrt{n}}\right), \widehat{\lambda_n}\left(1+\frac{q_{\alpha/2}}{\sqrt{n}}\right)\right].$$

Thus for this data set,

$$\begin{aligned}\mathcal{I}_{plug-in} &= \left[\frac{1}{7.8}\left(1-\frac{q_{0.05}}{\sqrt{64}}\right), \frac{1}{7.8}\left(1+\frac{q_{0.05}}{\sqrt{64}}\right)\right] \\ &\approx [0.1019, 0.1545].\end{aligned}$$

Since

$$\begin{aligned}|0.1064 - 0.1613| &= 0.0549 \\ |0.1019 - 0.1545| &= 0.0526,\end{aligned}$$

this implies that $\mathcal{I}_{plug-in}$ is the **narrower** confidence interval.

Finally,

$$\begin{aligned}((0.1064+0.1613)/2)*7.8 &= 1.04403\\ (0.1019+0.1545)/2)*7.8 &= 0.99996,\end{aligned}$$

so we see that $\mathcal{I}_{plug-in}$ is **centered** about $\widehat{\lambda}_n$, while $\mathcal{I}_{solve}$ is not. Alternatively, one can see directly from the formulas that $\mathcal{I}_{plug-in}$ is always centered about $\widehat{\lambda}_n$ whereas $\mathcal{I}_{solve}$ is **not** in general.

Submit You have used 3 of 5 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Foundation of Inference:Lecture 5: Delta Method and Confidence Intervals / 10. Confidence Interval for an Exponential Statistical Model

Add a Post

Show all posts ▼ by recent activity ▼

? Argument for the CI to not be of the form... 1

Learn About Verified Certificates