8. Exercise: Confidence interval interpretation

Exercise: Confidence interval interpretation

4/4 points (graded)

Every day, I try to estimate an unknown parameter using a fresh data set. I look at the data and then I use some formulas to calculate a 70% confidence interval, $[\widehat{\Theta}^-, \widehat{\Theta}^+]$, based on the day's data.

Are the following statements accurate?

Over the next 100 days, I expect that the unknown parameter will be inside the confidence interval about 70 times.

Yes ▼ ✔ **Answer:** Yes

If today's confidence interval is $[0.41, 0.47]$, there is probability 70% that the unknown parameter is inside this confidence interval.

No ▼ ✔ **Answer:** No

Out of 100 days on which the confidence interval happens to be $[0.41, 0.47]$, I expect that the unknown parameter will be inside the confidence interval about 70 times.

No ▼ ✔ **Answer:** No

不是某个特定的区间
No randomness here

Today, I decided to use a Bayesian approach, by viewing the unknown parameter, denoted by Θ, as a continuous random variable and assuming a prior PDF for Θ. I observe a specific value x, calculate the posterior $f_{\Theta|X}(\cdot \mid x)$, and find out that

$$\int_{0.41}^{0.47} f_{\Theta|X}(\theta \mid x)\, d\theta = 0.70.$$

Am I allowed to say that there is probability 70% that the unknown parameter is inside the (Bayesian) confidence interval $[0.41, 0.47]$?

Yes ▼ ✔ **Answer:** Yes

Solution:

The first statement is true. The confidence interval is a random interval and has probability 0.70 of capturing the true value of the unknown parameter. Using the frequency interpretation of probabilities, we expect about 70 successful captures.

The second statement is false. The value of the parameter is not random. Conditional on the confidence interval being $[0.41, 0.47]$, the event “the unknown parameter is inside the confidence interval" does not involve anything random, and so its probability cannot be 0.70.

The third statement may appear to be closer to the first one rather than the second one. However, the same explanation as for the second statement applies.

The fourth case involves the conceptually different setting of Bayesian inference. Here, Θ is a random variable, and

$$0.70 = \int_{0.41}^{0.47} f_{\Theta|X}(\theta \mid x)\, d\theta = \mathbf{P}\big(\Theta \in [0.41, 0.47] \mid X = x\big),$$

is indeed the (conditional) probability that Θ belongs to the interval $[0.41, 0.47]$.

提交 You have used 1 of 1 attempt