MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.s077 — Spring 2018
Recitation 2 — Friday 2/16

Topic 1. Confidence intervals for multiple parameters.
Confidence intervals have been defined for individual estimated parameters. If we are dealing with multiple parameters, what do these confidence intervals tell us for the parameter vector? This is the subject of this exercise.

Suppose that we have a method for constucting 95%-confidence intervals, denoted $I_1(X)$ and $I_2(X)$, for parameters θ_1 and θ_2. Here, X are the data, and our notation is chosen to emphasize the fact that the confidence intervals are determined by the data. In particular, $\mathbb{P}\big(\theta \in I_i(X)\big) = 0.95$, for $i = 1, 2$.

(a) We would like to construct something like a confidence interval (a "confidence set," or "confidence rectangle") for the vector (θ_1, θ_2). We woiuld like to make a statement of the form

$$\mathbb{P}\big(\theta_1 \in I_1(X) \text{ and } \theta_2 \in I_2(X)\big) \geq \beta.$$

What is the largest β for which such a statement is guaranteed to be true?

(b) Show how to construct a 90% confidence interval for $\theta_1 + \theta_2$. Your construction should be in terms of the end points of $I_1(X)$ and $I_2(X)$.

(c) Suppose now that $\hat{\Theta}_1 = \theta_1 + W_1$ and $\hat{\Theta}_2 = \theta_2 + W_2$, where W_1 and W_2 are independent normals, with common variance σ^2. (This situation might arise if we sample from two distinct populations and compute the sample mean of each population.)

(i) Construct a 90% confidence interval for $\theta_1 + \theta_2$, using the method from part (b). (Your answer should be involve $\hat{\Theta}_1$, $\hat{\Theta}_2$, n and σ.)

(ii) Construct a 90% confidence interval for $\theta_1 + \theta_2$, based on the estimator $\hat{\Theta}_1 + \hat{\Theta}_2$. Which one is narrower?

Solution

(a) Let, A_i be the event that $\{\theta_1 \in I_i(X)\}$, for, $i = 1, 2$. We are given that, $\mathbb{P}(A_i) \geq 0.95$, and we are asked to find the largest β for which $\mathbb{P}(A_1 \cap A_2) \geq \beta$. We have,

$$\begin{aligned}
\mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cup A_2) \\
&\geq 1.9 - \mathbb{P}(A_1 \cup A_2) \\
&\geq 0.9,
\end{aligned}$$

where, the first equality is a well-known identity, the first inequality follows from our assumptions on $\mathbb{P}(A_i)$, and the last inequality follows from the fact that $\mathbb{P}(E) \leq 1$, for any event E. Hence, $\beta = 0.9$ works out.
In order to see that this cannot be improved, simply consider any example where the probability mass of, $\mathbb{P}(A_1 \setminus A_2) = \mathbb{P}(A_2 \setminus A_1) = 0.05$, and, $\mathbb{P}(A_1 \cap A_2) = 0.9$ (here, $A_1 \setminus A_2$ stands for $A_1 \cap A_2^c$).

(b) We are guided from the previous part. Let, $I_i(X) = [A_i, B_i]$, for $i = 1, 2$, with $A_i < B_i$ (here, we are using uppercase letters to indicate that such confidence intervals are also random). Notice that

$$\theta_i \in [A_i, B_i] \implies A_1 + B_2 \leq \theta_1 + \theta_2 \leq B_1 + B_2.$$

Hence,

$$\mathbb{P}(\theta_1 + \theta_2 \in [A_1 + A_2, B_1 + B_2]) \geq \mathbb{P}(\theta_1 \in [A_1, B_1] \cap \theta_2 \in [A_2, B_2]) \geq 0.9.$$

(c) (i) Assume that, we have computed $\hat{\Theta}_1$, using n data points. Next, using the fact that the random variable

$$\frac{\sqrt{n}(\hat{\Theta}_1 - \theta_1)}{\sigma}$$

is approximately standard normal, we are asking for a t, for which we have 95% confidence. For this, recall that,

$$\mathbb{P}(|Z| \leq 1.96) = 0.95,$$

hence, we arrive at,

$$\mathbb{P}\left(\left|\frac{\sqrt{n}(\hat{\Theta}_1 - \theta_1)}{\sigma}\right| \leq 1.96\right) = 0.95,$$

in other words, the desired confidence interval is,

$$\left[\widehat{\Theta}_1 - 1.96 \cdot \frac{\sigma}{\sqrt{n}}\ ,\ \widehat{\Theta}_1 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Similarly, we have the following interval for θ_2:

$$\left[\widehat{\Theta}_2 - 1.96 \cdot \frac{\sigma}{\sqrt{n}}\ ,\ \widehat{\Theta}_2 + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Therefore, using the idea of the previous part, we arrive at the following confidence interval for $\theta_1 + \theta_2$:

$$\left[\widehat{\Theta}_1 + \widehat{\Theta}_2 - 3.92 \cdot \frac{\sigma}{\sqrt{n}} \ , \ \widehat{\Theta}_1 + \widehat{\Theta}_2 + 3.92 \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Furthermore, if we don't know the variance, we can replace the standard deviation term above, with $\widehat{\sigma}$, where, $\widehat{\sigma}$ is the approximate standard deviation, and is given by

$$\widehat{\sigma} = \sqrt{\sum_{i=1}^{n} \left(X_i - \frac{1}{n} \sum_{i=1}^{n} X_i \right)^2}.$$

(ii) For this part, note that the observation model is now,

$$\widehat{\Theta}_1 + \widehat{\Theta}_2 = \theta_1 + \theta_2 + W_1 + W_2,$$

where, $W_1 + W_2$ is a zero mean normal random variable, with variance $2\sigma^2$. Now, we use the fact that,

$$\frac{\sqrt{n}(\widehat{\Theta}_1 + \widehat{\Theta}_2 - (\theta_1 + \theta_2))}{\sigma\sqrt{2}}$$

is approximately standard normal, we arrive at the following confidence bound:

$$\mathbb{P}\left(\left| \frac{\sqrt{n}(\widehat{\Theta}_1 + \widehat{\Theta}_2 - (\theta_1 + \theta_2))}{\sigma\sqrt{2}} \right| \leq 1.65 \right) = 0.90.$$

In other words, the desired confidence interval is,

$$\left[\widehat{\Theta}_1 + \widehat{\Theta}_2 - 1.65 \cdot \frac{\sigma\sqrt{2}}{\sqrt{n}} \ , \ \widehat{\Theta}_1 + \widehat{\Theta}_2 + 1.65 \cdot \frac{\sigma\sqrt{2}}{\sqrt{n}}\right]$$

This confidence interval is narrower, and therefore more informative, than the one obtained in part (c)(ii).