Lsn23

Clark

Admin

Recall that a Confidence Interval is found by:

While this is helpful in most cases, in some cases we want to find a one-sided confidence interval. To do this we would instead:

Today we're going to look at some specific confidence intervals that come up quite a bit in statistics (for more see MA376!). Recall that under appropriate conditions, we can use a large sample approximation for the distribution of our pivotal quantity:

$$\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} \sim N(0,1)$$

Manipulation of this quantity gives a familiar form of a 95% CI for θ of

$$(\hat{\theta}-1.96\sigma_{\hat{\theta}}, \hat{\theta}+1.96\sigma_{\hat{\theta}})$$

Where we obtained the 1.96 from:

Or a 1-sided Confidence interval (Upper Bound) can be found from

$$(-\infty, \hat{\theta} + 1.65\sigma_{\hat{\theta}})$$

Note the difference in the constant value.

A quick picture:

If we want to use this to form a 95% CI for $E[Y] \equiv \mu$ we can use Table 8.1 and make the appropriate replacements:

$$\bar{Y} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{Y} + 1.96\frac{\sigma}{\sqrt{n}}$$

Sometimes Statisticians get asked to tell a researcher how many subjects they need in their study if they want their results to be precise to some measure. For instance, a state wildlife service wants to estimate the mean number of days that each licensed hunter actually hunts and they want their estimation error to be within 2 days.

Since they are estimating $\bar{Y}$ where Y_i is the number of days hunter i actually hunted, what they are saying (assuming they are forming 95% Confidence intervals is they want

$$1.96\frac{\sigma}{\sqrt{n}} = 2$$

Obviously to answer this we need σ. Our book claims that we may have an approximation of σ from previous studies. I am rarely so lucky. So perhaps we do this for a variety of σ values and give them a range of particpants. For instance if $\sigma = 10$ then

However if $\sigma = 20$ we would have:

In order to think of possible values of σ it may be helpful to consider what σ is. It is how much we deviate, on average, from the mean. So if your Y_i values are all close together, then σ is low, if not, then σ is high.

We can do the same thing if we estimate p by using $\hat{p} = \frac{Y}{N}$. Using Table 8.1 our asymptotic CI becomes:

Let's work through this with exercise 8.73 on page 424.

If we aren't so lucky that we can rely on asymptotics, we still have hope in forming a pivotal quantity for μ. However, in this instance we need our data, $Y_i \sim N(\mu, \sigma)$. If this is the case, then we can use

$$\sqrt{n}\frac{\bar{Y} - \mu}{S}$$

as our pivotal Quantity. What is the distribution of this?

Similarly, if we might be looking to build a confidence interval for, say, $\mu_1 - \mu_2$. Or the difference between two populations. If n is sufficiently large we can make use of the CLT and state that $\bar{Y}_1 \sim N(\mu_1, \sigma_1^2/n)$ and $\bar{Y}_2 \sim N(\mu_2, \sigma_2^2/n)$, so we then know the distribution of $\bar{Y}_1 - \bar{Y}_2$ is

So we can just replace σ_1 by S_1 and σ_2 by S_2 (why?) and use the following as a pivotal quantity:

If n isn't sufficiently large that the CLT kicks in we can still form a confidence interval for $\mu_1 - \mu_2$ if $Y_{1,i} \sim N(\mu_1, \sigma)$ and $Y_{2,j} \sim N(\mu_2, \sigma)$. However it takes a bit of work to figure out the distribution a pivotal quantity.

In order to figure this out, note that we are making the assumption that $\sigma_1 = \sigma_2 = \sigma$. Therefore, we can use both sets of data, $Y_{1,i}$ and $Y_{2,j}$, to estimate σ^2. One way to do this is if $i \in \{1, \cdots, n\}$ and $j \in \{1, \cdots, m\}$, we can pool our sample variance by

Then, it turns out we can find a pivotal quantity as

which can be used to find a confidence interval for $\mu_1 - \mu_2$.

In summary, we've talked about the following: