Lesson 33: How confident are you? What do you mean?





Although they look like trees, they are actually colonies of genetically identical animals.

M. Singh

Scenario:

Marine scientists say 10% of world's reef systems have been destroyed.

→ pollution, global warming, outright destruction, ocean acidification etc. may take out 70% of the reefs in 40 yrs

7 1 lab sampled sea fans at 19 locations along Yucatan Peninsula.

→ Found 54/104 sea fans infected with a disease in the sample collected at depth of 40ft at Las Redes Reef in Akumal, Mexico.

We care about the health of coral reef communities throughout the carribean. What can this study tell us about the prevalence of disease among sea fans?

This is one sample only

Sample proportion
$$\hat{p} = \frac{54}{104} = 51.9\%$$

Due to sampling variability another sample of 104 sea fans taken at same time might not give the same p-hat How can we find the actual population proportion? What can we even say about P with all the sampling variability?

**Imagine how sample proportions might vary from sample to sample

Central Limit Theorem assures us that our model for sampling distributions is approximately normal given certain assumptions are satisfied.

Is p = 51.9%? No! $\hat{p} = 51.9\%$ This is just one estimate





It says about 68% of all samples of 104 sea fans will have \hat{p} within 1SD of p

95% of all samples, \hat{p} will be within 2SD's of p



We are 95% confident that between 42% and 61.7% of Las Redes Sea fans are infected.

One - proportion Z-interval

95% confidence

We mean that 95% of all samples of this size will produce confidence intervals that capture the true proportion.

"confident"

Our uncertainty is about whether the particular sample we have at hand is one of the successful ones or one of 5% that fail to produce an interval that captures the true value.



End with having the whole class explain