Layman's Guide to Confidence Intervals and Statistical Significance

Confidence intervals and statistical significance are specific terms that are used in scientific literature. Although these terms might sound very mathematical, a reader doesn't have to be a mathematician to understand them and interpret the results.

Building your Confidence in Confidence Intervals

A confidence interval is simply a range within which a true value is likely to fall, based on the data used in the analysis.

In many instances, public health data are collected from a subset (sample) of a population, rather than from the entire population due to logistical and resource constraints. For example, often, an immunization rate is an estimate of immunization coverage, rather than the true rate, which you could calculate if you knew every vaccine given to every person in a group. The estimate may, in fact, be very close to the true rate, or it may be higher or lower than the true rate. To address this uncertainly, estimates can be presented with a confidence interval. A 95% confidence interval tells us that we can be 95% certain that the true rate lies somewhere between the lower and upper limits of the confidence interval.

For instance, look at this example:

Group	Pneumococcal Vaccination Rate for Adults ≥65 Years of Age	95% Confidence Interval
Less than a High School Education	56.8%	51.0% – 62.6%
College Graduate	71.2%	68.1% – 74.2%

Data Source: Behavioral Risk Factor Surveillance System/ CDC website, California Data, 2011

For older adults with less than a high school education, the rate of vaccination was estimated to be 56.8%. This means that among the sample of older adults included in this study, 56.8% of those who had less than a high school education, were vaccinated. The 95% confidence interval shows the range in which the "true" vaccination rate will fall 95% of the time based on the data utilized. In this instance, we can be 95% confident that the true rate lies between 51.0% and 62.6%. A narrower range, or confidence interval, would indicate that the sample data used in the study had less variation, thus strengthening our confidence that the estimate is close to the true vaccination rate in the population.

The Significance of Statistical Significance

Statistical significance, which is often referred to as just significance, means that a difference that is observed in study results is unlikely to be due to chance alone.

For example, statistical tests may reveal that the subjects who took a new allergy medicine had fewer runny noses than the subjects who took an older medication or that older people were more likely to have cancer than young people. If results are statistically significant, we can reasonably conclude that the difference in rates of runny noses and cancer are not just due to chance.

When looking at the above example of pneumococcal vaccination, the researcher asked the question "Are rates different for older adults based on their education level?" There are many different types of tests that you can use to test for statistical significance, based on the data that you have and what you want to ask. The above table represents one type of statistical testing that can be used to answer this question.

- In this case, because the confidence intervals do not overlap, we know that the rates are statistically different from each other. Therefore, we can say with confidence that adults over the age of 65 who graduated college received a pneumococcal vaccination more often than adults over the age of 65 with less than a high school education.
- However, if the two intervals DO overlap, the two results are not statistically different. Therefore, we CANNOT say with confidence that adults over the age of 65 who graduated college have higher vaccination rates than those with less than a high school education. Any differences in rates might be due to chance.
- AND, if there is no listed confidence interval, we CANNOT say that the percentages are statistically different. We CANNOT say that the vaccination rates of college graduates and those with less than a high school education are truly different.

The ABCs of the p Value

You can also determine whether there are statistically significant differences, by looking at the p value.

When interpreting data, a p value cut-off is set to signify whether differences are statistically significant. Generally in scientific literature, you see p value cut-offs set at 0.05 (95% confidence), but this is not an absolute cut-off. If the reported p value is 0.05 or less, there is a 5% or less chance that the effect you see is due to chance alone. So, generally if the p value is 0.05 or less, it is considered statistically significant. If the p value, is larger than 0.05, the results are not statistically different.

However, remember that with a 95% confidence interval, there is still a one in twenty chance that this difference is not "real" and was only observed due to chance. In other words, if you took a sample of the data 100 times, 5 times you would get a result that was outside the report range (95% confidence interval), by luck of the draw. Remember, increasing the sample size, or collecting data from a larger proportion of the population, has the effect of decreasing the p-value. This is because larger samples tend to be more representative of the target population. That's all. That's the big secret.

Resources for Learning More

- **What are Confidence Intervals and P-values?**
  www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_are_Conf_Inter.pdf
- **Annenberg Learning Statistics Instructional Videos and Learning Guides**
  www.learner.org/resources/series65.html?pop=yes&pid=3158
- **How to read a paper: Statistics for the non-statistician I: Different types of data need different statistical tests, BMJ**
  www.bmj.com/content/315/7104/364
- **How to read a paper: Statistics for the non-statistician. II: "Significant" relations and their pitfalls, BMJ**
  www.bmj.com/content/315/7105/422

Contact Us

To learn more or to join the CIC Disparities Workgroup, please contact Julia Heinzerling, MPH at jheinzerling@ph.lacounty.gov or 213-351-7800.