Testing Hypothesis and the p-value

Estimating a parameter and confidence intervals

* If we have the relevant data for the entire population, we can simply calculate the parameter.
* If the population is very large, a statistic based on a random sample can be a reasonable estimate of an unknown parameter in the population.
* The value of any statistic depends on the sample, and the sample is based on random draws. A question that arises:
  + “How different could this estimate have been, if the sample had come out differently?”
* Bootstrap: a method to generate confidence intervals that uses resampling (with replacement) of the sample.
  + Why does it work?
    - By the law of averages, the distribution of the original sample is likely to resemble the population, and the distributions of all the “resamples” are likely to resemble the original sample. So the distributions of all the resamples are likely to resemble the population as well.
* CI is an interval which is expected to typically contain the parameter being estimated. It refers to the probability that a population parameter will fall between two set values.

How are p-value and confidence interval related?

* **P-value** is a measure of strength. It indicates the probability of obtaining the observed result, if the null hypothesis is true. If the p-value is less than the chosen significance level (such as 0.05 or 0.01), we reject the null hypothesis in favor of the alternative hypothesis.
* A **confidence interval** is a range of values that is likely to contain the true value of a parameter. It provides information about the precision of our estimate.
* If a confidence interval contains the null hypothesis value, we fail to reject the null hypothesis
* If a confidence interval does not contain the null hypothesis value, we reject the null hypothesis

Our approximate 95% confidence interval for the average age in the population goes from 26.9 years to 27.6 years. Suppose someone wants to test the following hypotheses:

**Null hypothesis:** The average age in the population is 30 years.

**Alternative hypothesis:** The average age in the population is not 30 years.

Then, if you were using the 5% cutoff for the p-value, you would reject the null hypothesis. This is because 30 is not in the 95% confidence interval for the population average. At the 5% level of significance, 30 is not a plausible value for the population average.

This use of confidence intervals is the result of a *duality* between confidence intervals and tests: if you are testing whether or not the population mean is a particular value *x*, and you use the 5% cutoff for the p-value, then you will reject the null hypothesis if *x* is not in your 95% confidence interval for the mean.

This can be established by statistical theory. In practice, it just boils down to checking whether or not the value specified in the null hypothesis lies in the confidence interval.

If you were using the 1% cutoff for the P-value, you would have to check if the value specified in the null hypothesis lies in a 99% confidence interval for the population mean.