

NR140 Group Work z-tests

On blackboard you have the global land and sea temperature anomalies from 1880 to 2016 as recorded by NASA (Source: NASA Goddard Institute for Space Studies, "Global Temperature Anomalies in 0.01 C" at **Global-mean monthly, seasonal, and annual means**, 1880-present, updated through most recent month) <https://data.giss.nasa.gov/gistemp/>

Typically, NASA climatologists use 1981-2010 as a **baseline climate "population"**. We have a column that labels each year as either "historical (pre 1981)", "baseline population (1981-2010)" or "Anthropocene (2011 on)".

Our task will be to **compare the mean Historical and Anthropocene Annual climate anomalies to this Baseline "population" annual anomaly using a one-sample z-test.**

I will get you started with information for some of the steps we need to consider when conducting an inferential test. Your job is to complete the rest of these steps (specifically answer the questions in red below).

Let's start in R but keep in mind that you can run a one sample z-test in excel using the =ztest(array, pop_mean, pop_stdev) function.

*Note that we will be using the **Annual Anom J-D** data column (although climate data for all of the individual months and seasons is also included).*

Start with a comparison of **historical** era anomalies to the **baseline population** following our steps for any inferential test:

1. State the null hypotheses:
H0a: There is no difference between the **mean of historical era annual anomalies** and the **baseline population mean**.
2. Set alpha level for the test
P<0.05
3. Choose the appropriate test
We are **comparing a sample mean (historical era) to a specified population (baseline population) mean**. This should be a one **sample z-test** as long as data is normally distributed.

4. Let's check for normality
 - a. Run a Shapiro Wilks test on the **annual anom J-D data column for the historical** era (since this is the sample we want to test against the population, the sample is what we must test for normality)...so be sure to **by the era** categories.
 - b. **Report the p-value for the goodness of fit test for the historical data only**
 - c. **Is your historical annual anom normally distributed?**
5. Calculate the test statistic (in this case the one sample z-test). In order to run a one sample z-test we first need to know the mean of the sample we want to test and the mean and standard deviation of the population we are testing our sample against. Since all of our data is stacked in one column, we will need to separate our eras using the filter function from tidyverse.
 - a. Report the **mean of the baseline population era annual anom J-D**. This will be your hypothesized population mean that you test.
 - b. Report the **standard deviation of the baseline population era annual anom J-D**. This will be your hypothesized population standard deviation that you test
 - c. From the "z.test" report of the **historical era sample** that you want to compare to this baseline population, **report the obtained (calculated) z-test statistic**.
 - d. **Does the historical era mean fit the baseline data?**

Your final task is to repeat this whole process but this time comparing the few years we have in the "Anthropocene" era to the baseline population mean and write a summary paragraph

Specifically, test to see if the Anthropocene Era has been significantly higher than the baseline population of climate anomalies.

To give you some perspective on why we would ask such a question: if we find that our recent Anthropocene era climate differs significantly from the baseline climate (our hypothesized population mean), it infers that our recent climate comes from a different population (i.e. the nature of climate itself has changed significantly and the climate we live in now does not come from the same population of temperatures that we consider our baseline "normal". It would mean we are in a new Era of climate conditions.

I won't walk you through this step by step, but let you walk through the analysis on your own, ending by summarizing it all in a concise, clear paragraph.

6. Summarize: **Write a concise on paragraph summary of the Anthropocene sample mean comparison to the baseline population.**

