Is the observed global temperature increase over 1949-2008 statistically significant?

An illustration of the Monte Carlo (or bootstrap) method

1. a) Imagine that your friend claims to have developed a special technique for rolling dice such that she can make three red dice always have a greater sum than three white dice. Before doing so, she lets you roll the dice as much as you want to convince yourself that they are non-loaded dice. Your friend then provides one, and only one, demonstration of her technique. There appears to be no obvious difference between the way she rolls dice and the way you roll dice. Your friend rolls the dice, and the sum of the white dice is ten and the sum of the red dice is eleven. Do you believe that your friend can really make the red dice always have a greater sum than the white dice? Why or why not?

   b) Imagine that your friend instead rolled the dice and the sum of the white dice was three and the sum of the red dice was eighteen. She refuses to provide a second demonstration. In this case, do you believe that your friend can really make the red dice always have a greater sum than the white dice? If your answer is different from before, why is that?

   c) Would you believe your friend if she rolled the dice and produced a sum of eight for the white dice and a sum of twelve for the red dice? Why would your answer be the same or different from before?

   d) Considering that your friend will give only one demonstration, can you think of another way to test whether your friend really has a special dice rolling technique?

2. a) Download and open the Excel file for the dice activity from the course website.

   b) Roll the three white dice and three red dice 100 times. Enter the white dice sum and the red dice sum in the specified columns. The red-white column automatically calculates their difference. Check the mean of the 100 red-white difference values. Is it exactly zero? Would you expect it to be zero? Why or why not?

   c) Look at the histogram chart of the red-white difference values. What kind of shape does it have? Why do you think it has that shape? Check the standard deviation of the 100 red-white difference values. The standard deviation is the typical magnitude of deviation between a single red-white difference and the mean red-white difference. Looking at the histogram chart, what percentage of values are within the standard deviation?

   d) Sort the red-white differences from most negative to most positive (Hint: highlight the white, red, and difference columns, then go to Data → Sort → Sort by "Difference" Ascending). What is the 95th sorted value? Is it larger than 1? Is it larger than 4? Is it larger than 15? What is the approximate percentage chance that a random roll of three red dice and three white dice will produce a red-white difference greater than the 95th sorted value? Do you want to change any of your answers to part 1?

3. a) In part 1, your friend claimed to have a special technique for rolling dice such that the sum of three red dice was always greater than the sum of three white dice. One way to test this hypothesis would be through repeated demonstrations, but she would only do it once. It is often the case in climate science that repeated demonstrations are not possible, so we need to develop another method for testing hypotheses. In particular, the biggest question regarding a single demonstration is whether the supposed phenomenon happened by coincidence (i.e., your friend had no special technique and rolled a red sum greater than the white sum merely by "luck") or by real effect (i.e., your friend really did have a special technique). How could we test a hypothesis in this way?

   b) Formally, we will call "rolling a greater red sum by luck" the *null hypothesis* because this is the outcome we would ordinarily expect and thus is the best starting point. The *alternative hypothesis* is "rolling a greater red sum by technique". We will reject the alternative hypothesis if we can show that the result of your friend's single roll was likely to happen by chance in the context of the null hypothesis. We will reject the null hypothesis if we can show that the dice roll result was unlikely to happen by chance. The common standard used in science is 5% (or 95%); that is, the alternative hypothesis is rejected if the result is more than 5% likely to have happened by chance, and the null hypothesis is rejected if the result is less than 5% likely to have happened by chance. Note that meeting the 95% confidence level (i.e., confident that it is 95% likely the effect is real) does not infallibly prove the alternative hypothesis since it is still 5% likely to have happened by chance (a false positive or type I error). Similarly, not meeting the 95% confidence level does not infallibly disprove the alternative hypothesis since there still may be a real effect (a false negative, or type II error). Explain your activity in part II in terms of null hypothesis and alternative hypothesis. What difference between the sum of red dice and white dice would your friend need to obtain in one roll to reject the null hypothesis at 95%? At 90% or 99% (other confidence levels used in science)?

4. a) You have a friend who is a global warming skeptic. To convince him that climate change is real, you show him how global mean temperature has increased during the past six decades. Each value in the table below is the approximate temperature difference between one decade and the decade of 1949-1958 (e.g., the decade of 1999-2008 was 0.55°C warmer than the decade of 1949-1958).

Decade 1	Decade 2	Decade 3	Decade 4	Decade 5	Decade 6
1949-1958	1959-1968	1969-1978	1979-1988	1989-1998	1999-2008
0.00°C	0.05°C	0.05°C	0.20°C	0.35°C	0.55°C

You point out that the sum of decades 4-6 is 1.0°C larger than the sum of decades 1-3 (this corresponds to a 0.33°C average temperature increase). Your friend counters that Earth's temperature randomly fluctuates from one decade to another, and there is no reason to believe that the observed increase from 1949-1978 to 1978-2008 occurred other than by chance. How could you demonstrate the warming is likely to be associated with the increase in CO_2 rather than a random fluctuation? What would be the null and alternative hypotheses?

b) The most appropriate null hypothesis is that the observed warming from decades 1-3 to decades 4-6 is due to natural random fluctuations (i.e., not related to anthropogenic or other outside effects). The alternative hypothesis is that the difference is not random but instead is related to outside effects. The null hypothesis would be rejected if we could demonstrate that it is less than 5% likely that the observed warming could happen by chance. How could we do this?

c) In order to determine whether the observed warming could happen by chance, we need to know what sort of temperature fluctuations could possibly occur. The simplest assumption is that the set of possible temperature fluctuations is the same set as those that have been observed (e.g., the temperature value from one decade to another is simply a random selection from the six values in the table). How could we carry out an experiment?

d) Roll the three white dice and three red dice 100 times. Convert the number on each die to a particular temperature value from the table. Sum the three white temperature numbers and the three red temperature numbers and enter the values in the specified columns. The red-white column automatically calculates their difference. Check the mean of the 100 red-white difference values. Is it exactly zero? Would you expect it to be zero? Why or why not?

e) Look at the histogram chart of the red-white difference values. What kind of shape does it have? Why do you think it has that shape? Check the standard deviation of the 100 red-white difference values. Looking at the histogram chart, what percentage of values are within the standard deviation?

f) Sort the red-white differences from most negative to most positive. What is the percentage chance that the observed increase in sum from decades 1-3 to decades 4-6 would be larger than 1.0°C? What confidence do you have that the observed increase is not merely random? Is the null hypothesis rejected or not?