

PHSX815_Project1: Hypothesis Test for the rate of Sunny Days per Year

Ghadah Alkhalifah

February 2023

1 Introduction

Many parts of the world are expected to see more sunny days as a result of global warming. This is because more heat held in the atmosphere, making the sky clearer and less covered in clouds.

This paper is organized as follows: Sec. 2 explains the hypotheses we are testing to see if we are expecting more sunny days in the next year or not. A description of the computer simulation developed to simulate these possibilities is provided in Sec. 3, with an analysis of the outputs included in Sec. 4. Finally, conclusions are presented in Sec. 5.

2 Hypotheses to test if we have a significant change in the sunny days rate next year

In order to determine whether there is evidence of a dramatic increase of the rate of sunny days, I set the following hypotheses:

1. The null hypothesis (H_o): The rate of sunny days is 122 days a year.
2. The alternative hypothesis(H_a): The rate of sunny days is more than 122 days a year.

3 Code and Experimental Simulation

Let's look at some simulated data: First, I assume the weather condition can be categorized as sunny, rainy, and cloudy days [1]. The probability for each category is set to be equal to check the possibility of having more sunny days than other weather conditions in one year (i.e. 365 days). To do that, I use a random number generator to generate a categorical distribution [2] [3]. Then, since the sunny days (events) occurred independently with a rate every year [4], I use a Poisson generator from Numpy library to draw a sample of random discrete probabilities of the expected sunny days' rate a year. Also, I use the Poisson probability mass function pmf() method [5].

To test the hypotheses, a student t-test was used [6] to compare the generated sample mean (or rate) to the mean specified by the null hypothesis [7]. The significance level was set to 10% and 5% .

4 Analysis

For the first simulation that I use categorical distribution for weather conditions with equal probabilities, the number of sunny days counts for one year was 124 days (i.e. higher than rainy and cloudy days). The bar chart in Fig.1 shows the counts of weather conditions.

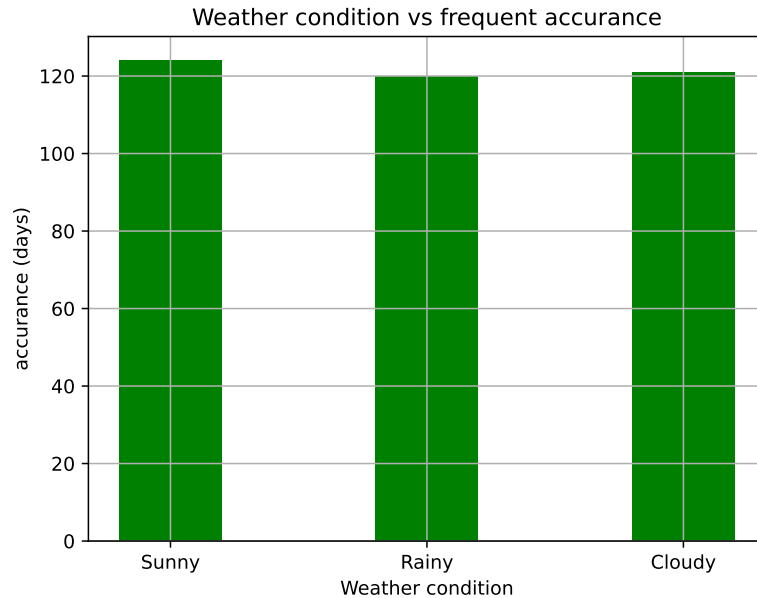


Figure 1: The outcome of categorical distribution of weather conditions with equal probabilities.

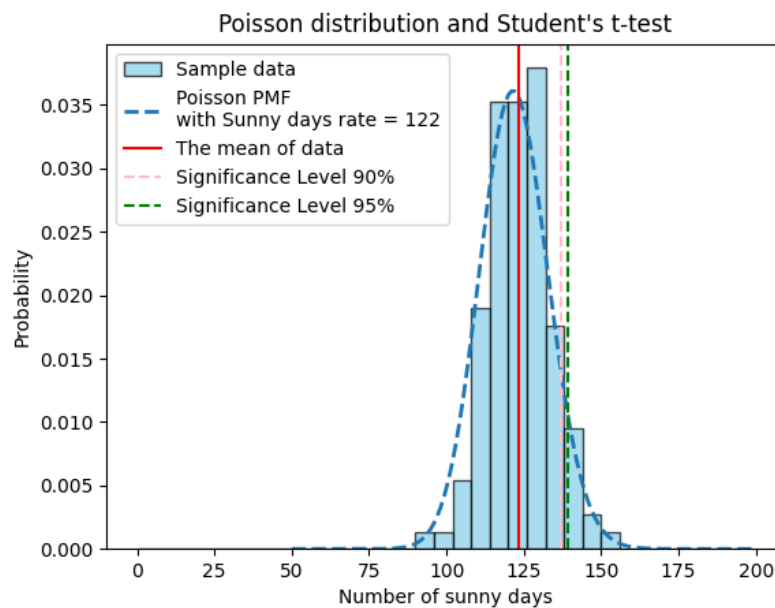


Figure 2: The Poisson PMF plot and student t-test for rate= 122 days

I set the sample size 123 in Poisson generator and the rate of sunny days as 122 days an in the null hypothesis. So, from Figure.2 I can see that the generated sample mean of the observed sunny days per year was 123.18. The t-statistic was calculated to be 1.32, and the p-value was 0.095. The p-value is less than the significance level of 0.1, so we reject the null hypothesis. As I decrease the significance level (alpha) to 0.05, we fail to reject the null hypothesis. There is not enough evidence to support the alternative hypothesis that the mean number of sunny days per year is greater than 122.

5 Conclusion

Based on the results of this analysis, I cannot conclude that the expected rate of sunny days per year is greater than 122 days. Further testing may be needed to obtain sufficient evidence to support the alternative hypothesis.

Here are the references

- [1] *Categorical distribution and Indicator Function*
<https://www.youtube.com/watch?v=421uW9aZHio>. Accessed: 2023-02-06.
- [2] *Read a file line by line in Python*
<https://www.geeksforgeeks.org/read-a-file-line-by-line-in-python/>. Accessed: 2023-02-06.
- [3] *PHSX815-Week2* https://github.com/crogan/PHSX815_Week2/tree/master/python. Accessed: 2023-02-01.
- [4] *Poisson Distribution and Poisson Process in Python* <https://pyshark.com/poisson-distribution-and-poisson-process-in-python/#poisson-distribution-example>. Accessed: 2023-02-02, Author: Misha.
- [5] *How to Create a Poisson Probability Mass Function Plot in Python?*
<https://www.geeksforgeeks.org/how-to-create-a-poisson-probability-mass-function-plot-in-python/>. Accessed: 2023-02-07, Last Updated : 13 Jan, 2022.
- [6] *How To Perform A Student's T-Test In Python*
<https://predictivehacks.com/how-to-perform-a-students-t-test-in-python/>. Accessed: 2023-02-07.
- [7] *Using SciPy T.ppf to get p-value* <https://stats.stackexchange.com/questions/271258/using-scipy-t-ppf-to-get-p-value>. Accessed: 2023-02-07.