wwwwwAnalytics for Observational Data (IT142IU)

Lab 5: Sampling and Hypothesis testing

## Objectives

* Understanding hypothesis testing
* Apply hypothesis testing techniques to some datasets.
* Dataset sources:
  + <https://www.kaggle.com/doj/federal-firearm-licensees>
  + <https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data>
* Programming languages: Python/Java
* Ref: <https://www.kaggle.com/residentmario/hypothesis-testing-with-firearm-licensees> [1]

## Tasks

***Part 1. Follow the instructions in [1] and describe the steps of hypothesis testing in your report.***

|  |  |
| --- | --- |
| Dataset | Firearm Licensees |
| Study question | Number of gun sold per shop is 2.75 |
| H­o | H0 : n =2.75 |
| H1 | Ha: n =/= 2.75 |
| Calculate a test statistic |  |
| Calculate a p-value |
| Make a decision and interpret your conclusions | Our P value is 0.72, it mean there is 72% that H0 is likely to occurs. Since P-value is 0.72 > 0.05( 5% risk when rejected H0), we fail to rejected H0, and there is strong evidence said that n = 2.75 (almost 3) |

***Part 2. Apply hypothesis testing to predict the land average temperature.***

|  |  |
| --- | --- |
| **Questions** | **Answers** |
| Dataset | GlobalTemperatures.csv |
| Study question | Is the land average temperature 8.37? |
| H­o | H0: m =8.37 |
| H1 | H1: m =/= 8.37 |
| Choose the sample dataset | Sample size = 1000 |
| Calculate a test statistic |  |
| Calculate a p-value |
| Make a decision and interpret your conclusions | Assume our alpha value is 5% and our P-value is 0.85 => Fail to reject H0 |

|  |  |
| --- | --- |
| **Questions** | **Answers** |
| Dataset | GlobalLandTemperaturesByCity.csv |
| Study question | Did Latitude affect Average Temperature? |
| H­o | H0: Latitude do affect Average Temperature? |
| H1 | Ha: Latitude do not affect Average Temperature |
| Choose the sample dataset | N= 100 |
| Calculate a test statistic |
| Calculate a p-value |
| Make a decision and interpret your conclusions | 1. Value suggest that there is extremely small percent that H0 is likely to happen= > rejected H0 |