AIIP 2018 Statistics Assessment Report

Student: Philippus Rudolf Viljoen

# Study Design

**Sampling discussion:**

Given the known differences in certain variables, the questionnaire was designed in such a way to control these variations. Where bias could not be controlled, sources were randomly selected. In this case households were randomly selected. The study employed multi-stage stratified random sampling.

**Bias**

Many procedures in the sampling design aimed at ensuring reliable and unbiased data production. However, it is impossible to not have bias in a study like this.

Answers are mostly based on people knowledge, memory and opinion. Answers could also be filtered or the interviewer’s presence could affect the response in an unwanted way.

Open questions were used in section VII, Adaptation options. Farmers were asked about their experience and perception of climate changes, their principle adaptation strategies and their constraints for implementing adjustments.

It was not financially feasible of practically possible to do multiple visit surveys to reduce memory and recall biases because of the wide geographic spread of the sample.

The study design was as appropriate as practically possible considering the scale and wide geographic spread of the survey. It certainly includes some self-selection bias in the results. This will have to be considered during further statistical analysis.

This was an observational study which means that relationships shown to be statistically significant can only be considered associative and not necessary causatory.

# Key Questions

Based on the dataset, I would be interested in answering the following two questions;

1. Does the climate change have a greater negative impact on income generated from farming in South Africa or Zimbabwe?
2. Is there correlation between the farmers education level and their ability to adapt to climate change?

# Statistical Analysis Plan

Let’s use question 1 in above to come up with a statistical analysis plan to come to a inference.

My hypothesis tests will be;

null hypothesis, H0: Climate change has no or a smaller negative impact on income generated from farming in South Africa than Zimbabwe

alternative hypothesis, Ha: Climate change has a greater negative impact on income generated from farming in South Africa than Zimbabwe

To test this, I will use a 1 sample t-test. I chose this test because of the following reasons:

* I want to compare the difference between a population mean and a hypothesized value.
* Y is continuous
* X is discrete
* I need to focus on the centering of data

To determine statistical significance, I will use a tool such as Minitab to calculate the descriptive statistics of my dataset and confirm the correct hypothesis test to run. I will then run the test and calculate the P-value. If the probability (P-value) is less than alpha (I’ll use 0.05) then I will conclude that the outcome is evident of an unusual difference from my null hypothesis. The observed value is thus statistically significant.

# Results Interpretation

If a significant effect resulted I would calculate the confidence interval. This will give me a range of values that my point estimate may fall within. The interval estimate can also help me conclude the true population parameter. I can also calculate the confidence level to check whether it is likely that the confidence intervals will contain the population parameter.

If the results were insignificant, I will review my analysis plan to find a more suitable method of analysis, and also consider a more appropriate sampling strategy.