Statistics Summative Assessment 11th February 2019

Submitted by Rose Funja - rfunja19@aiip.alueducation.com

Question1

Qualitative and Quantitative Variables.

Qualitative refers to data that deals with descriptions.It can further be divided to nominal eg ***gender variable*** and ordinal data eg ***farmtype variable.***

Quantitative refers to numerical data and can be divided into Interval and ratio data example ***hhsize*** variable measuring the size of the household and ***education*** variable as ratio type that counts the number of years spent in school with zero meaning absence of formal education

Question2

The study design used two methods 1. stratified and 2.multi-stage sampling combined called multi-stage stratified random sampling. Firstly eleven countries were selected to represent the four sub-regions of Africa East, West, North and Southern Africa. Secondly districts selection was done within selected countries from each sub-region to capture representative farms across diverse agro-climatic conditions within each country. Stage three sampling involved selection of villages and finally in each district, surveys were conducted of clustered farms randomly.

The above method used is appropriate since the application of the survey was to inform about the importance of each system for a country's and Africa as a whole agricultural production and its ability to cope with short- and long-term climate changes or extreme weather events.

Convenience bias sampling was clustered in villages to reduce the cost of administering the survey

Non-response bias :Respondent cannot remember details e.g. on labour distribution among different household members or on details for time periods in the past.

Question3 Research questions

- Is subsidies (crop and input) the most available form of financing for smallholder farmers?

- What is the relationship between farmer's ability access to extension services and farm yield?

Question4

Choice:Research Question 2

Null Hypotheses: There is a significant difference in yield for farmers who access extension services

Alternative Hypotheses: There is no significant difference in yield for farmers who access extension services

Test: t- test for independent variables because data is normal, data has equal variance and the observations in the dataset are independent of each other

Determining Significance:

T-statistic, sample size and P-value (or confidence interval) are the key measures for interpreting this test.

P-value is the probability of observing an outcome at least as extreme as the sample statistic, given the null hypothesis is true. If this probability is less than 0.05 or 5%, the outcome is concluded to be evidence of an unusual difference from the null hypothesis. The observed value is then said to be statistically significant.

Question5

1) Reject the null hypothesis

No difference in yield for farmers who have access to extension services

2) Keep the null hypothesis

Theres is difference in yield for farmers who have access to extension services

Question6

1. Explore relationship between variables: **scatterplots**

yield against **access** to extension services

2. Compare data sets:**Histogram**:- for common crops in the four geographical locations: **Average crop type yield** per **country**
3. Explore the constituent parts of a whole group:**pie chart:-Gender, age and geographical** distribution of the farmowners

Question7

Is it important that data on fertilizer use, pesticide use and the irrigated area was compared to national average values from the FAO and the World Bank.This is so as to reduce the non-sampling errors that are introduced due to inappropriate analysis conclusions and false information provided by respondents. Since the aim is to make conclusion about the population it is important that the samples are representative and standardized.