Stat 231 - Statistics **Spring 2017**

Lecture 2: May 3rd, 2017

Lecturer: Suryapratim Banerjee *Notes By: Harsh Mistry*

2.1 Terminology

Definition 2.1 *Population is a collection of certain properties we are interested in.*

Definition 2.2 *Inductive : Small → Big*

Definition 2.3 *Deductive : Big → Small*

Definition 2.4 *Empirical Studies are studies where data is collected wither through observation or through experimentation. Repeated studies will often given different results*

Definition 2.5 *Experimental Studies are studies where some of the variables are controlled by the data collector. For example in a medical study you may control who receives actual treatment and who receives a placebo treatment.*

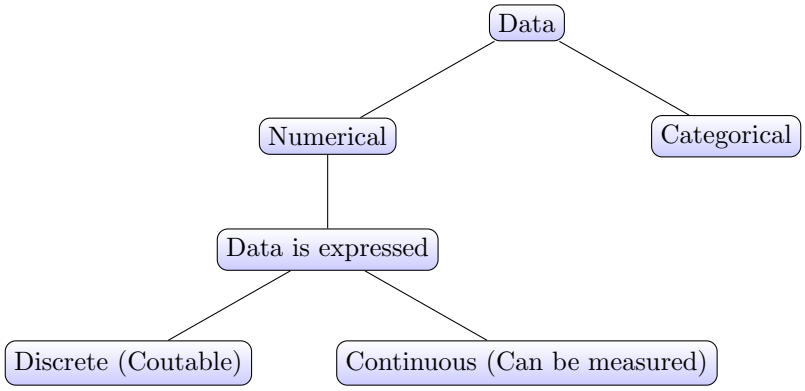
Definition 2.6 *Observational Studies are studies where the data collector has no control the variables. For example, a study between alcohol consumed and GPA would have no variables that can be controlled.*

Definition 2.7 *Unit : Each member of the population*

Definition 2.8 *A Variate is the property of the units we are interested in*

Definition 2.9 *A Attribute is a function of the variables.* $f(y_1, \ldots, y_n)$

2.2 Types of Data



Definition 2.10 *Categorical Data is non-numerical data. For example, Yes/No/Unsure are example of none numerical data.*

Definition 2.11 *Binary Data Two categories*

Definition 2.12 *Ordinal Data is Categorical data which has an underlying order.*

Definition 2.13 *Non-Ordinal Data is Categorical data which has no underlying order.*

Definition 2.14 *Coding is the process of converting categorical data into numerical data. For example :*

$$\begin{cases} \textbf{\textit{No}} \ \to 0 \\ \textbf{\textit{Maybe}} \ \to 1 \\ \textbf{\textit{Yes}} \ \to 2 \end{cases}$$