Lec 1-

Let's do a survey. Who has an Phone? Stordard notation for a "datum" and the same X2=0, X3=1, X4=1, X5=0 2800 X6=1,1,1,0,0,1,1,1,1,0,0,1,1,0 finst respondent 1-20 in our "sample" 12 (203) , 8 (003) Do we believe this survey is a sample of n=20 elements from a superset called the "population"? If we do, this is called the "Regulation model sampling assumption" It so, what is that population? - All people of Farth - All people of America + All people in Q.C. - All college students Is this sample representative of the population? This is typical. Given a sumple, assume a population model, then identify the representative popularions This tappens in data science all the time. In classical statistics, this goes the apposite direction. You begin by defining the population clearly i.e everyone who has als over 60, clear population. Then sample of elements from that populations

Population has size N. You have some idea of what N is.

If pop= all Americans => N= 330 million.

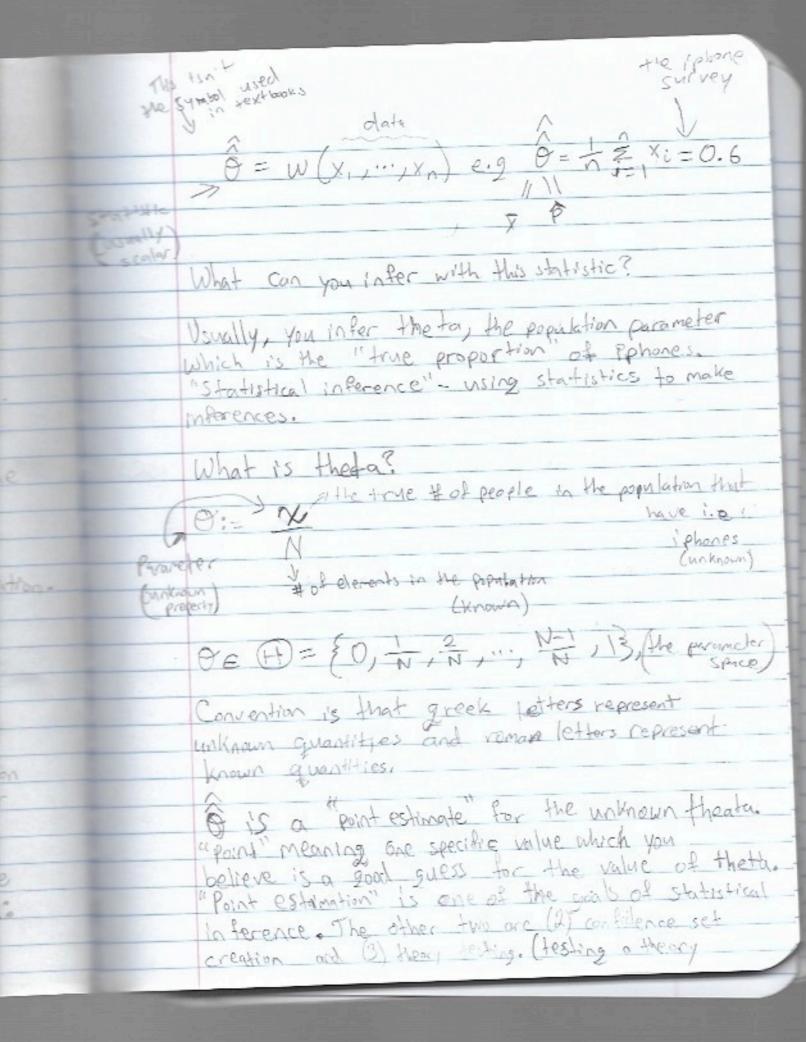
Can we learn about the population from the sample?

Yes. This 95 called "inference" We use the sample to "infer" properties about the population.
Usually the properties are parameters of the

"Infer" means to make an educations guess from specific things to universal properties.

A synonym is "induction". The opposite is deducation which is universal > particular. You can never be sure your inference is correct.

How is inference done with data? You generate "Statistics" which are functions of the datas



about a specific value of theta at a accertainty level" alpha). Let's sangle one element from the population. And do one survey. How should this element be chosen of I want a "representative" sample? = Randomly but specifically, uniformly meaning every element has oprobability of 1/2 of thing chosen. That's called a "simple corden sample" (St i.e. imagine a hat with names in it. So, What is the probability that X1=1? P(X = x, -1) = 2 = 0 - specific value the C.V. the survey realization a rapid is and