module_prob2_lecture10

Fri, 1/21 12:45PM 16:35

SUMMARY KEYWORDS

party, electorate, probability, support, young voters, voters, constitute, survey, age, calculation, similarly, older, fraction, extrapolate, votes, conditional probabilities, marginal probability, important, population, total

SPEAKERS

Sumon Majumdar

S Sumon Majumdar 00:04

Since the last clip, I introduced you to this data based survey of 1000 voters about the support for three different possible political parties, party A party B and party C. And the question here is that if we were to use the data from the survey, to extrapolate it and try and form predictions about which party is more likely to win the election, then how would you go about doing that? And this is particularly interesting in this model, because see different groups support different parties, right. So among the younger voters, that is under 30 years, most of the support the biggest supporters for party A, among the middle age voters, it's for party C, and for the older ones, its support for party A, but only slightly over policy. So now, if we use this, this, this two way frequency table, right, so it's a two way table, because we have two characteristics. One is parties. And the other thing here is the ages. Right? So if we look at the total number of voters under 30. So if we add all of these up, right, 180, 120, 100, that's 400, a- the middle age voters 110 121 70. That's also an out of 400 and older voters at 50 and 70. That's 200, right. And so that adds up to our 1000. And similarly, you can figure out what's the total number of people who support party A, party B and party C, right? So for example, for party C, it's 100 plus 170? That's 270 plus 70. That's 340. For this, it's 290. And for party A, 180 plus 110. That's 290 plus 180. That's 370. Now, so what's the sur- In the survey, right? If you ask the question, what's the probability that a random person in the survey is under the age 30 years and supports party? So what we are looking at suppose we want to look at the probability of a party supports party a and is under 30. So that's young, right? So what we'll have to look at is the number. So that's this cell here. So the probability of that will be 180, divided by the total number of people in the survey. So that's 180, over 1000. So that'll be point one, eight, right, and so on. Now, you can also ask the question about the marginal rate, what's the support for party a among the under 30s? Right? What's the support for party B among the under 30s? What's the support for party C among the under 30s? So these are conditional probabilities. So if you're looking at the support of A, among the younger voters, so there's a total of 400 younger voters, right, so that goes in the denominator. What's the support for party A, that's this, so that's 180 over 400. And that'll turn out to be .44 Similarly, if you're looking at what's the support for party B among the young voters, again, there are 400. And if you look at them, right, so the 120 support them, so that's 120 over 400. And that will be .3. Similarly, for party C, among young voters, there are 400 And how many of them support so that's 100 over 400, that's going to be .2. So these are

the conditional probabilities of support for the different parties among the young. Now, you can ask the same question On the middle aged voters as well, right, so if you look at that there's a total of 400 Middle Age voters. And what's the support for party A? So that's 110 divided by 400. All right, for for party B, it's 120 divided by 400. For party see, it's 170 divided by four. And similarly, you can ask for the older voters also, the older there are 200. Right? So if you look at the support for party A among older voters, that will be at over 200. Okay. So, and so, right, so, this will be .4. And similarly, we can calculate the others as well. So, these are not conditional probabilities. Now, the question is, how can we use this? So, to ask the following sort of question, right, is that, can we take the results from the survey, and an extrapolated in order to get an idea of the outcome among the whole population? So, if you think about it, right, so, what will be important in that is in the whole population, what fraction of the whole population is young voters? What fraction of the whole population is middle aged voters, and what fraction is older voters? Because these three groups support different parties, and different. So, the the composition of the electorate, in terms of the ages will be important in determining which party is more likely to win. Right. So let's try to do that here. So, my first question is that suppose each group, right each age group, constitutes 1/3 of the electorate. So suppose there is equal so 1/3 of the electorate is young 1/3 is middle it and 1/3 is older. Right? Which party is likely to? So what we're interested in is, what's the probability for party A? What's the probability for party B? What's the probability for party C? And this will give us what fraction of the total votes each of these parties are likely to win? So how do we do that? So first thing, see if we look at party a, so there's 1/3 of the population is young. Right? So so it's the probability of young times the probability of a given young, so that's point four, five. Similarly, what's the probability of mid that's 1/3 times the probability of a given it? That's point 275. Plus another 1/3? The probability of holder and what's the probability of a given older that's point four. And if you add all of these up and do the calculation, this comes out to be point 375. What about party B? Right, so for party B, again, we'll do the same thing with probability wanted its young, right? And the support given young for party B is point three than 1/3. Its middle aged support for party be given middle age, that's point three, and 1/3 is older and support there is .25. And if you do the calculation here, this turns out to be .283. So this is the marginal probability for B of support for B, this is the marginal probability of support for a. Now what's the marginal probability of support for party C? Again, we do the same sort of calculation 1/3 is young support among young for party C is .251. One third, it's mid and support among them is .425, and 1/3 is .35, right? Okay. And if you do this calculation, this turns out to be .342. And Now if we're going to compare among these three, right, so if you look at a voter from this electric, then what's the probability of support? The highest is for party A it's 37.5%. Or in other words, 37.5% of the electorate are likely to support party A based on the results of this survey. Right. So based on this results of the survey, if each age group constitutes 1/3 of the electorate, then based on the survey, party A is likely to get the highest number of votes. And you can sense that this is very important, like what's very important in this calculation is how is the age distribution of the electorate? Right here, I've considered where there is an equal number of young, middle and older voters. That may not be the case, the age composition could be different in depending on like, it's a different province, or a different country, right? Or whether you're looking at one constituency versus another constituency. Right. So let's consider another case. What if this is a different constituency, where the young voters constitute 25% of the electorate, the middle aged voters constitute 50%. Right? And older, they constitute the rest, which is 25%. So it here now, what's the marginal probability of A, B and C? So we do exactly the same calculation here, but with these new fractions, right, so young voters now constitute only one quarter of the of the electorate, right. So the support among them for party a is point four, five. Now the middle age voters, they constitute half the electorate, for among them for party A is .275. And among the older voters who constitute one quarter, the support is .4. So if you add these three up, that's .35, right. So the support for party A in the whole electorate

based on the survey is going to be 35%. Similarly, for party B, let's do one quarter times, this is point three, half, which is the middle age voters times this is point three, another one quarter is older. And that's point two, five, right? And if you add all these up, that's .2875. So party B is definitely going to be getting lesser votes than party A. Right? What about party C? Remember, party C, had a strong following among middle aged voters right. Now? If they constitute a bigger chunk of the electorate? Does it mean that party C has a higher probability of winning? Right? So let's calculate that. Right? So one, four, of the voters a Yeah. And among parties for party C, their support is .25, half our middle, and their support is point four to five and a quarter are old. And support for party C is .35. And if you do this calculation, this is .3625. And now see, if you compare between the probability of A, B and C, right, this is the probability of support for party A. In the electorate, this is that for party B in the electorate, and this is for party C, and we see that the support for party C is the highest right? It's slightly higher than that of party A. Right so based on the results of the survey, if this is the age composition of the electorate, right then party C is more likely to so this is how conditional probabilities marginal probabilities, they play an important role often in real life decision making. Right? So what we have done is taken a situation very plausible situation, like do a, do a survey of voters, right? Figure out what's the probable support for each of the parties and extrapolate that results from there for the whole election. Right, based on what you know about the demography mix of the electorate, in terms of their age, composition, and political parties to all the this all the time, the break up the electorate based on levels of education, gender, age, income, and then try and figure out what's the support for their party relative to the other parties in each of this subgroups of the population. And then they try and use that to figure out what may be the support and what may be the total number of votes that they may get in the electorate as a whole. So let's stop this clip here. And in the next clip, of take the same idea of conditional probability, but do something interesting with it, which is reverse the order of conditioning and introduce you to the idea of the famous Bayes rule.