Transcript

Live Video – Examples (A and B) – Part 2

01. 00:01 / 00:05 - Let's go back to this example A and B that we started with earlier. So suppose eight
02. 00:05 / 00:10 - percent of US adults twenty years or older have diabetes again we could say that as the
03. 00:10 / 00:17 - prevalence of diabetes is 8 percent. These are the questions that we asked. If we select
04. 00:17 / 00:22 - n people at random from US adults, how many do we expect to have diabetes? What is the
05. 00:22 / 00:26 - probability that none of them will have diabetes? Or at least one of them will? That we can
06. 00:26 / 00:32 - answer now. Distribution and how variable is that distribution. Only the question in
07. 00:32 / 00:37 - bold we are going to look at right now. The others are later. What if I picked three people
08. 00:37 / 00:44 - at random from all US adults, what is the probability that none of them will have diabetes?
09. 00:44 / 00:51 - So if I pick the first person what's the probability that they will not be diabetic? It was 8 percent.
10. 00:51 / 00:58 - That do? So what percent don't? 92. So the probability that the first person does not
11. 00:59 / 01:06 - have diabetes is 0.92. Right? Second time I reach in what's the probability that the
12. 01:06 / 01:13 - person does not have diabetes? 0.92. Why? Because I'm drawing from every person in the
13. 01:13 / 01:18 - whole United States all US adults so if I reduce one million people by one, it makes
14. 01:18 / 01:24 - zero difference in my calculation in this scenario. Right? But if I was fixed in a small
15. 01:24 / 01:29 - group, now I have to worry which is the next example B. And in the third time it's still
16. 01:29 / 01:35 - going to be 0.92. So that's my answer for that. Independent, because it's a large population
17. 01:35 / 01:39 - as long as we're truly randomly selected. It's not true if I go pick three people from
18. 01:39 / 01:44 - the same family. Right? Then you might have either genetic or social reasons that you
19. 01:44 / 01:49 - might all be diabetic or not be diabetic. So again the probability that none are diabetic
20. 01:49 / 01:56 - is 0.92 times 0.92 times 0.92 or 0.92 to the 3rd which is about 78 percent. And then the
21. 01:59 / 02:06 - probability that at least one is diabetic is kind of little trick. If it's not true,
22. 02:06 / 02:13 - what's the opposite of having at least 1 person being diabetic if it's not true that at least
23. 02:14 / 02:20 - one person is diabetic how many people must have been diabetic. In this possibility we,
24. 02:20 / 02:26 - we, pick three people. Right? So my only choices are there could be no people that are diabetic,
25. 02:26 / 02:31 - there could be one person who is diabetic, there could be two people who are diabetic,
26. 02:31 / 02:37 - or all three people could be diabetic. So at least one would be one two or three people
27. 02:37 / 02:44 - are diabetic. So what's the opposite? None. It's always the case, the probability that
28. 02:44 / 02:51 - at least one something is always one minus the probability none whatever that thing is.
29. 02:53 / 02:57 - So the probability that at least one person is diabetic is always one minus the probability
30. 02:57 / 03:04 - that none diabetic, none are diabetic is easy to calculate we already did it before 0. 7787,
31. 03:04 / 03:11 - one minus that 0.2213. And then, let's look at the example B. Suppose in a group of 60
32. 03:12 / 03:17 - subjects 5 have diabetes we already answered the first question a long time ago. If we select
33. 03:17 / 03:23 - one person at random what is the probability the person is diabetic well there's five diabetics
34. 03:23 / 03:29 - inside 60 total people. So 5 out of 60 or about eight percent 8.33 percent. But now
35. 03:29 / 03:34 - let's start with two. If we select two subjects at random what is the probability that both
36. 03:34 / 03:41 - are diabetic? So the first time I reach in how many people are diabetic. out of how many
37. 03:42 / 03:49 - people total? And now the second time I reach in, I know the first person had to be diabetic,
38. 03:49 / 03:56 - so how many diabetics are left? Out of how many total people. Very good. So that's dependent
39. 03:56 / 04:01 - scenario we're really using a multiplication all here. We said the probability the first
40. 04:01 / 04:08 - person was diabetic times the probably the second person is diabetic given the first
41. 04:09 / 04:16 - person was. So, I don't think I wrote that notation out here but it is right now on the
42. 04:16 / 04:23 - one we did earlier from the data. So 0.0056 and then same deal here. We could do a none.
43. 04:23 / 04:28 - I didn't do an at least one but I could. But none are diabetic, well, first time I reach
44. 04:28 / 04:35 - in 55 out of 60 are not diabetic I know the first person was not diabetic, so the second
45. 04:35 / 04:38 - time I reach in there's only fifty four of them left. I can mix it. I could say was the
46. 04:38 / 04:45 - probably the first is diabetic the second is not. It's just a matter of keeping track
47. 04:45 / 04:50 - of what numbers are changing in which numbers are not Changing. So maybe you can think about
48. 04:50 / 04:57 - that one. I know that there's only three possibilities, either both, none, or one of each. And one

49. 04:58 / 05:03 - of each can happen in two ways the first one could be diabetic the second one not or vice
50. 05:03 / 05:07 - versa the first one not and the second one is. So there are two other probabilities to
51. 05:07 / 05:13 - calculate. And then later we'll be able to answer if we select two subjects at random
52. 05:13 / 05:16 - what's the expected number with diabetes.