

Transcript

Live Video – Basic Probability Using Logic

01. 00:00 / 00:04 - What we're going to do for the first questions is we're going to pick one person from our
02. 00:04 / 00:09 - sample. We haven't seen the data organized yet by the groups and yes/no. But we will
03. 00:09 / 00:14 - see that in a second. We are going to ask these following questions. Suppose one individual
04. 00:14 / 00:20 - is selected from our sample at random and consider the following questions. What is
05. 00:20 / 00:26 - the probability that the individual has these calcium oxalate crystals in their urine? That's
06. 00:26 / 00:30 - the first one. What is the probability that the individual has a calcium concentration
07. 00:30 / 00:35 - of five or more? That's the second one. So we're actually going to call those two events
08. 00:35 / 00:40 - A and B. And then, these are just combinations of those events. What is the probability that
09. 00:40 / 00:45 - the individual has these crystals present and has a calcium concentration of five or
10. 00:45 / 00:50 - more? And then, we'll say the or what is the probability that the individual has calcium
11. 00:50 / 00:55 - oxalate crystals present or has a calcium concentration of five or more. So, this is
12. 00:55 / 01:02 - like A, B, A and B, A or B. Those are just the basic common questions that you could
13. 01:03 / 01:10 - be asked. This is very similar to what you will see happening to you on quizzes and tests.
14. 01:10 / 01:15 - All of these four probability questions are going to relate to the entire sample meaning.
15. 01:15 / 01:19 - Our denominator is going to be the total number of people in our dataset end of story; I don't
16. 01:19 / 01:24 - even have to think about that, I know my bottom number once I look at the data. And here we
17. 01:24 / 01:31 - have our, table. What we can see here is the on the left we have each of the groups we
18. 01:32 / 01:38 - created from calcium concentration. It's between 0 and up to but not including two, between
19. 01:38 / 01:45 - two up to but not including five, between five up to not including 8, and 8 or more.
20. 01:45 / 01:49 - So we can see how many people that were no's in each of these groups and how many people
21. 01:49 / 01:54 - that were yes's in each of these groups. We can see the row totals and we can see the
22. 01:54 / 02:01 - column totals and this ought to sound, everything we do here ought to sound very similar to
23. 02:01 / 02:07 - the questions that were asked related to C-C; a categorical predictor and a categorical
24. 02:07 / 02:12 - outcome in the notes. Last time we would have called it a percent. Right? We would have
25. 02:12 / 02:17 - said what percent have calcium oxalate crystals present? And could have answered that question.
26. 02:17 / 02:21 - Right? So now we're just going to change the terminology. And we're going to say what's
27. 02:21 / 02:27 - the probability that a person has calcium oxalate crystals present? So here's our first
28. 02:27 / 02:32 - question; what is the probability that the individual has calcium oxalate crystals present?
29. 02:32 / 02:38 - Again, we will denote this event by A. That is for my convenience, it is unfortunate that
30. 02:38 / 02:43 - that is probably not your convenience because it means you have to translate back and forth
31. 02:43 / 02:50 - a lot .So A is that we do indeed have these calcium crystals present. We could call this
32. 02:50 / 02:55 - the prevalence of calcium oxalate crystals in our sample, but we would have to be very
33. 02:55 / 03:01 - careful about calling this an estimate for the population until we knew what this population
34. 03:01 / 03:06 - was. So again it might be the estimate if I'm talking about every person who is coming
35. 03:06 / 03:10 - to the urologist with the problem. But I'm going to guess this is not representative
36. 03:10 / 03:14 - of the whole population and the reason is let's look at what that number works out to
37. 03:14 / 03:21 - be. So again at 73 total, of those 31 said yes or were yes, so I get 31 out of 73, 42.5%.
38. 03:24 / 03:29 - So my guess is that a high number for the overall population of the whole United States.
39. 03:29 / 03:35 - I could be wrong; I don't know anything about calcium oxalate crystals. P stands for probability
40. 03:35 / 03:40 - we just shorten that notation into P parentheses whatever you want to write, any could be p
41. 03:40 / 03:47 - of A or if it's something that's kind of short you could write something like p male, you
42. 03:47 / 03:51 - know you can use the words of the problem, or you can use a letter that you create, or
43. 03:51 / 03:56 - we create . It is up to you sometimes but sometimes you have to translate ours. This
44. 03:56 / 04:02 - would be hard to fit in the notation. So I want to come up with A or B to keep it simple
45. 04:02 / 04:06 - in this in this case because calcium oxalate crystals present, that's a little bit big
46. 04:06 / 04:11 - to put my formulas. So here's the second basic probability. What is the probability that
47. 04:11 / 04:17 - the individual has a calcium concentration of 5 or more? This is a little harder than
48. 04:17 / 04:24 - that but not much. How many total people have a calcium concentration of five or more? 16

49. 04:24 / 04:31 - plus 8, right? Out of 73. So there you go total 73 we have to add these to row totals
50. 04:34 / 04:41 - 16, 8 to get 0.329. So we're going to call that the probability of B which represents
51. 04:43 / 04:50 - the probability that the individual has a calcium concentration of five or more. So
52. 04:50 / 04:56 - what is the probability that the individual has calcium oxalate crystals present and has
53. 04:56 / 05:03 - a calcium concentration of five or more? So now the answer is eleven plus seven, because
54. 05:03 / 05:08 - there are eleven people that were five to seven and yes and seven people who were eight
55. 05:08 / 05:15 - or more and yes. We were looking really at the intersection of these rows for 5 to 7.99
56. 05:16 / 05:21 - and eight or more the intersection between those rows and column of Yes's and we only
57. 05:21 / 05:26 - want to count people who are literally in both. You must have a calcium concentration
58. 05:26 / 05:33 - of five or more and have these crystals present. So there's eleven plus seven out of 73 or
59. 05:34 / 05:41 - 0.247. And we will be pulling these probabilities back out later because we going to use them
60. 05:41 / 05:44 - to convince you of the rules. And we are going to say, hey look, didn't we just do this really
61. 05:44 / 05:49 - easy. So you if don't need a rule, just do it by counting. And I'm telling you that most
62. 05:49 / 05:53 - the time I'm going to give you the ability to do it this logical way because I'm just
63. 05:53 / 05:57 - looking to see do you understand the basic concepts. Let's go to the "or" which is a
64. 05:57 / 06:03 - little harder. What is the probability that the individual has these crystals present
65. 06:03 / 06:08 - or has a calcium concentration of five or more. We spend a little time talking about
66. 06:08 / 06:14 - what do we mean by or and give the example of mom with a kid in the store, you can have
67. 06:14 / 06:21 - toy A or toy B. The kid does not have the choice to pick both, so in our mind or is
68. 06:21 / 06:27 - like often this exclusive or. I have to pick one or the other that is not what happens
69. 06:27 / 06:34 - in probability. Or is the inclusive or, meaning it is either A or B or both. So that's just
70. 06:36 / 06:43 - something you've got to learn. So we want to count everybody who was either a yes or
71. 06:44 / 06:49 - five or more. So I want to count everybody is in these squares but I do not want to count
72. 06:49 / 06:55 - these people twice. And I've got to count everybody who's five or more which is, you're
73. 06:55 / 07:01 - right, 24. And I want to count everybody who's yes which is, you're right, 31. But if I add
74. 07:01 / 07:08 - 24 to 31, I've counted eighteen people twice, that's not allowed. Probability is you, only
75. 07:08 / 07:15 - get one shot at being counted in probability. So you can see I just did the really blunt
76. 07:15 / 07:20 - way five plus one plus eleven plus seven plus nine plus four just so you'd see where all
77. 07:20 / 07:25 - these numbers were coming from. But I can't count anybody twice. This is one case where
78. 07:25 / 07:29 - if you mess this up a lot, just go with the rule because it'll fix it for you. It will
79. 07:29 / 07:33 - subtract off those you would be counting twice. And it'll save you from making that mistake.
80. 07:33 / 07:39 - But if that makes sense, we really don't need a rule. We can just add them together ourselves,
81. 07:39 / 07:46 - being careful not to count anybody twice. So we get 37 out of 73 or a little more than
82. 07:46 / 07:47 - fifty percent.