

# Transcript

## Live Video – Conditional Probability

01. 00:00 / 00:06 - Now I am going to mess up your world. So everything we've done so far we've always been divided  
02. 00:06 / 00:11 - by the total, 73. I think the worst thing that can happen to you is that you slightly  
03. 00:11 / 00:16 - misread a question in the early stuff. But here's what I know happens from my experience  
04. 00:16 / 00:21 - as a longtime teacher of these kind of classes. Once I teach you the harder concept, you will  
05. 00:21 / 00:28 - second-guess yourself on the easy concept all day. So I'll give you the tools to know  
06. 00:28 / 00:33 - when is it easy and when is it not as easy and hopefully that'll help you not let yourself  
07. 00:33 / 00:39 - get stressed out or answer the question incorrectly. But conditional probability is a new idea,  
08. 00:39 / 00:45 - we are going to start with the logic and then we'll go to the formula. So again, so far  
09. 00:45 / 00:52 - we've divided by the total, but sometimes we have additional conditions that cause us  
10. 00:52 / 00:58 - to alter the denominator, if we're doing it through logic, of our probability calculations.  
11. 00:58 / 01:04 - So the denominator is the bottom of our probability calculation. And the question is what  
12. 01:04 / 01:10 - condition have I been told that makes me adjust who I'm really interested in. If I'm not interested  
13. 01:10 / 01:16 - in everybody then I'm dividing by less than the total. So here's the kind of question  
14. 01:16 / 01:23 - that we will be asking. Given the individual has calcium oxalate crystals present, what  
15. 01:24 / 01:31 - is the probability the individual's calcium concentration is five or more? So the condition  
16. 01:31 / 01:38 - I've been asked to satisfy is that the individual must have calcium oxalate crystals present  
17. 01:38 / 01:42 - and then I'm going to answer this question. So you can think about is almost like reducing  
18. 01:42 / 01:48 - your sample I now no longer care about the no's. They're gone. I only care about the  
19. 01:48 / 01:55 - yes's in this question. And it's that given word that's the clue. I can also use if, right,  
20. 01:55 / 02:01 - if is really the same thing those are really the only two words I use on a regular basis.  
21. 02:01 / 02:08 - I don't try to confuse you. But given or if are the two, keywords telling you: this is  
22. 02:09 / 02:15 - conditional probability. I need to be careful. If you don't see given or if, then you don't  
23. 02:15 / 02:20 - need to be as careful and try really hard not to let yourself be fooled into making  
24. 02:20 / 02:25 - it harder than it really is. That's the best advice I can give you. Again, this word conditional  
25. 02:25 / 02:31 - probability refers to the fact that we have some additional condition or restriction or  
26. 02:31 / 02:37 - some other information that we're going to be using in our probability calculation. So  
27. 02:37 / 02:40 - again we're looking straight from data, let's look at the question, the same question we  
28. 02:40 / 02:46 - just ask given that the individual has calcium oxalate crystals present, how many people  
29. 02:46 / 02:53 - do I have to work with? So I was read that given and find my denominator. So 31 is who  
30. 02:53 / 02:59 - I'm going to divide by. And I can only account people that are in that yes column. How many  
31. 02:59 / 03:06 - of those 31 have calcium concentration above five or more? So I add the eleven plus seven  
32. 03:07 / 03:14 - 18 out of no longer 73 because I don't want out everybody out of 31 because I have been  
33. 03:14 / 03:21 - given that the person must be in the Yes column. And that's it. So here I'm blacking out my  
34. 03:21 / 03:28 - no column, blacking out my total column, I only need the yes column. Of those 31 total;  
35. 03:30 / 03:37 - and of those eleven plus seven are five or more. So given gives me sort of a new denominator  
36. 03:38 / 03:43 - to work with that's based upon only those people who satisfy that condition. Now we  
37. 03:43 / 03:48 - haven't talked about this notation yet, so, but this notation if I want to read this probability  
38. 03:48 / 03:53 - notation let's start with the one that doesn't have the blue over it. So P parentheses B  
39. 03:53 / 04:00 - vertical line A: that says the probability of B happening given A happened or given A.  
40. 04:02 / 04:09 - So given is always at the back whatever's in the back is my new bottom, my new total,  
41. 04:09 / 04:16 - Whatever satisfies the given is what I divide by in my fraction. But see again, no formula  
42. 04:17 / 04:22 - needed, just logic. You need to understand the concept of conditional probability this  
43. 04:22 / 04:27 - is insanely important to statistics. Because again this is how we make those comparisons  
44. 04:27 / 04:33 - among group one this happened, among group two this happen are they the same or different.  
45. 04:33 / 04:38 - That's what we want to learn. So conditional probability, very important when we actually  
46. 04:38 / 04:44 - start to apply these things in the real world. We might say is the prevalence of diabetes  
47. 04:44 / 04:50 - the same among males and females? yes or no. That's this question. Those are both conditional  
48. 04:50 / 04:55 - probabilities. Given your male, what's the probability you have diabetes? Given you are

49. 04:55 / 04:59 - female, what's the probability you have diabetes? Are those two numbers the same in the  
50. 04:59 / 05:05 - population or they are clearly different in the population? But here is the formula, and the formula  
51. 05:05 / 05:10 - just is again, a way that you can calculate that probability if you know probabilities already  
52. 05:10 / 05:15 - and you don't have the big count kind of table. So back here to this problem we could use  
53. 05:15 / 05:21 - this rule, again using we've already found we had the probability of A is 0.425, the  
54. 05:21 / 05:28 - probability of B is 0.329 and the probability of A and B is 0.247. what we want to calculate  
55. 05:28 / 05:35 - is the probability of B given A. Which what did we get on the previous slide 0.581, and  
56. 05:35 / 05:41 - so if we plug in these numbers, amazingly, not so amazingly actually, we get the right  
57. 05:41 / 05:46 - answer. So if I want to find the probability of B given A which is the probability that  
58. 05:46 / 05:51 - given you have calcium oxalate crystals present that you have a high calcium concentration  
59. 05:51 / 05:58 - of five or more. So five or more given yes and I plug in my A and B which was 0.247,  
60. 06:00 / 06:06 - that's where it came from originally 18 over 73, but it's 0.247. And the probability of  
61. 06:06 / 06:11 - A is 0.425 and by dividing these two numbers I get exactly what I got on the previous slide  
62. 06:11 / 06:18 - 0.581. So these rules are just giving you again a truth, it's true that that works.  
63. 06:19 / 06:26 - Let's do another practice given that the individual does not have calcium oxalate crystals present,  
64. 06:27 / 06:32 - what is the probability that the individual's calcium concentration is five or more? So  
65. 06:32 / 06:38 - how many total do I have to work with if they do not have calcium oxalate crystals present?  
66. 06:38 / 06:45 - And of those how many have a concentration is five or more? So six out of forty-two.  
67. 06:47 / 06:54 - And that probability is 0.143. Now compare those two answers, very very different. Right?  
68. 06:57 / 07:04 - So this is the probability that you have a concentration of five or more among the yes  
69. 07:04 / 07:09 - group is fifty-eight percent; the probability you have a calcium concentration of five or  
70. 07:09 / 07:16 - more in the No group is only 14 percent. Very different, not independent clearly dependent,  
71. 07:17 / 07:22 - clearly the probability is different in those two groups. And I don't really need to test  
72. 07:22 / 07:27 - this to know that I believe it in the population whatever this population is. That's really  
73. 07:27 / 07:34 - what I am going to ask you A, B, A or B, A and B, A given B. You can definitely expect  
74. 07:37 / 07:42 - all five of those questions with data like this on your quizzes and tests. Now, that  
75. 07:42 / 07:47 - doesn't mean that all I'm going to do, but those will be there. You don't need me to  
76. 07:47 / 07:48 - tell you that again.