Lecture 10: Independence

Anup Rao

April 22, 2019

We discuss independent events and random variables.

Recall these basic facts:

Fact 1 (Chain Rule). *If $A_1, A_2, \ldots, A_n$ are events, then*

$$p(A_1 \cap A_2 \cap \ldots A_n)$$
$$= p(A_1) \cdot p(A_2|A_1) \cdot p(A_3|A_1 \cap A_2) \cdot \ldots \cdot p(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

In the past few lectures, we have been using the Chain rule over and over to compute various probabilities. The chain rule is even easier to use with *independent* events. Two events A, B are said to be independent if $p(A|B) = p(A)$. This can only happen when $p(B|A) = p(B)$, since by Bayes' rule, we have

$$p(B|A) = \frac{p(A|B) \cdot p(B)}{p(A)} = p(B).$$

Example: n coin tosses

Suppose you toss n coins independently. What is the probability that they all come up heads? On the one hand, we can calculate this directly as $1/2^n$, since there is only 1 outcome that is all heads, and 2^n possible outcomes in total. But we can also calculate this by writing H_i to be the event that the i'th coin is heads. Then the coin tosses are all mutually independent. Then the probability that they all come up heads is

$$\begin{aligned} p(H_1 \wedge H_2 \wedge \cdots \wedge H_n) &= p(H_1) \cdot p(H_2|H_1) \cdot \ldots \cdot p(H_n|H_1, \ldots, H_{n-1}) \\ &= p(H_1) \cdot p(H_2) \cdot \ldots \cdot p(H_n) \\ &= (1/2)^n. \end{aligned}$$

Limited Independence

Just because every pair of events in a collection of events is independent, that does not mean that all the events are mutually independent. Consider this example—pick a random string $x \in \{0,1\}^3$ subject to the constraint that the string has an even number of 1's. Let A_1, A_2, A_3 denote the events that the first, second and third coordinate are 1's. Then $p(A_1|A_2) = p(A_1) = 1/2$, and the same is true for every pair of random variables. However, if we calculate the probability of getting the string 111, this is

$$\begin{aligned} 0 &= p(A_1 \wedge A_2 \wedge A_3) \\ &= p(A_1) \cdot p(A_2|A_1) \cdot p(A_3|A_1, A_2) \\ &\neq p(A_1) \cdot p(A_2) \cdot p(A_3) \\ &= 1/8. \end{aligned}$$

Even though every pair is independent, $p(A_3|A_1, A_2)$ is actually 0 and not 1/2.

It is easy to generalize the same example—if you pick a uniformly random string from $\{0,1\}^n$ subject to the constraint that the number of 1's is even, then every $n-1$ of the bits will be independent and uniformly distributed. These kinds of examples mean that it is very hard to use data to test that a collection of variables is independent without sampling the entire probability space. In class, we discussed how this causes problems with evaluating DNA evidence in forensics.