Chris Piech
CS109

Independence

Independence is a big for machine learning. The joint probability of many events requires exponential amounts of data. By making accurate independence and conditional independence claims, computers can calculate accurate probabilities very fast, based on a lot less data.

Independence

Two events E and F are called independent if: P(EF) = P(E) P(F). Or, equivalently: P(E | F) = P(E). Otherwise, they are called dependent events

Three events E, F, and G independent if:

P(EFG) = P(E) P(F) P(G), and
P(EF) = P(E) P(F), and
P(EG) = P(E) P(G), and
P(FG) = P(F) P(G)

If events E and F and independent, E and F^C are also independent. Independence unfortunately is not transitive.

Generally n events E_1 , E_2 , ..., E_n are independent if for every subset with r elements (where $r \leq n$) it holds that: $P(E_a,E_b,\ldots E_r) = P(E_a)P(E_b)P(E_r)$. For example: the outcomes of n separate flips of a coin are all independent of one another. Each flip in this case is called a "trial" of the experiment.

Hash Map Example

Let's consider our friend the hash map. We are going to hash m strings (unequally) into a hash table with n buckets. Each string hashed is an independent trial, with probability p_i of getting hashed to bucket i. Calculate the probability of these three events:

A) E = *at least one string hashed to first bucket*
B) E = At least 1 of buckets 1 to k has ≥ 1 string hashed to it
C) E = Each of buckets 1 to k has ≥ 1 string hashed to it

Part A

Let F_i be the event that string i is not hashed into the first bucket. Note that all F_i are independent of one another. By mutual exclusion $P(F_i) = (p_2 + p_3 + \ldots + p_n)$

P(E)	$= 1 - P(E^C)$	Since $P(A) + (A^C) = 1$
	$= 1 - P(F_1 F_2 \ldots F_k)$	By the semantics of F_i
	$= 1 - P(F_1)P(F_2)\ldots P(F_m)$	Since the events are independent
	$= 1 - (p_2 + p_3 + \ldots + p_n)^m$	We calculate $P(F_i)$ by counting equally likely.

Part B

Let F_i be the event that at least one string is hashed into bucket i. Note that F_is are neither independent nor mutually exclusive.

P(E) $= P(F_1 \cup F_2 \cup \ldots \cup F_k)$
$= 1 - P([F_1 \cup F_2 \cup \ldots \cup F_k]^C)$ Since $P(A) + (A^C) = 1$
$= 1 - P(F_1^C F_1^C \ldots F_k^C)$ By DeMorgan's Law
$= 1 - (1 - p_1 - p_2 - \ldots - p_k)^m$ By thinking about the semantics

The last step is calculated by realizing that $P(F_1^C F_1^C \ldots F_k^C)$ is only satisfied by m independent hashes into buckets other than 1 through k.

Part C

Let F_i be same as in Part B.

P(E) $= P(F_1 F_2 \ldots F_k)$
$= 1 - P([F_1 F_2 \ldots F_k]^C)$ Since $P(A) + (A^C) = 1$
$= 1 - P(F_1^C \cup F_2^C \cup \ldots \cup F_k^C)$ By DeMorgan's Other Law
$= 1 - P\left(\bigcup_{i=1}^{k} F_i^c\right) = 1 - \sum_{r=1}^{k} (-1)^{(r+1)} \sum_{i_1 < \ldots < i_r} P(F_{i_1}{}^c F_{i_2}{}^c \ldots F_{i_r}{}^c)$ By Inclusion/Exclusion

Where $P(F_1^C F_1^C \ldots F_k^C) = (1 - p_1 - p_2 - \ldots - p_k)^m$ just like in the last problem.

Conditional Independence

Two events E and F are called conditionally independent given G, if
$P(E F \mid G) = P(E \mid G) P(F \mid G)$
Or, equivalently: $P(E \mid F G) = P(E \mid G)$

Conditional Breaking Independence

Let's say you see a lawn and it gets wet when it rains or when the sprinklers turn on. For this particular set of sprinklers, rain and sprinklers are independent events. In general knowing if the sprinklers went off does not tell you whether or not it rained. Now, you come outside and see the grass is wet. Your belief that it rained has gone up and your belief that the sprinklers went off has gone up. Both explain why there may be wet grass.

In this situation, given that the grass is wet, gaining the knowledge that the sprinklers went off would change your belief that it has rained. The sprinklers explain why the grass is wet, and so the alternate explanation becomes less likely. The two events (which were previously independent) are dependent when conditioned on the grass being wet.

Conditioning on an event E leads to dependence between previously independent events A and B when A and B are independent causes of E.