Notes 20

1 Independence

Definition 1.1 *An event A is independent of an event B if*

$$Pr(A|B) = Pr(A).$$

In other words, knowing that B occurs doesn't change the probability that A occurs.

For example, if we flip two coins, then the result of coin 1 is unrelated to that of coin two, thus the event "coin 1 is head" is independent of the event "coin two is head".

Independent events are particularly nice to work with. This is because there is a very simple rule to compute the probability of their intersection.

Theorem 1.2 *If A is independent of B, then*

$$Pr(A \cap B) = Pr(A) \cdot Pr(B)$$

Proof.

$$Pr(A \cap B) = Pr(B) \cdot Pr(A|B) = Pr(B) \cdot Pr(A)$$

■ Note that the product rule is symmetric in A and B; it is not by chance.

Theorem 1.3 *If A is independent of B, then B is independent of A. In other words, independence is a symmetric relation.*

Proof.

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

■

Thus it is appropriate to say that A and B are "independent of each other". This makes it possible to define mutual independence of more than two events.

Definition 1.4 *Events $A_1, A_2, \cdots, A_n$ are mutually independent, if for all i, $1 \le i \le n$ and for all J, $J \subseteq [1,n] - \{i\}$ we have*

$$Pr(A_i | \bigcap_{j \in J} A_j) = Pr(A_i)$$

In other words, a collection of events is independent, if each of them is is independent of the intersection of any others. Equivalently, the above events are mutually indenpendent, if $\Pi_{i=1}^{n} Pr(A_i) = Pr \cap_{i=1}^{n} (A_i)$.

For example, if we toss n coins and A_i is the event that the i-th coins is head, then these events are mutually independent, as $Pr(A_i) = 1/2$ and this is true regardless of the results of other coins.

The following example is a little bit tricky. We flip three fair coins. Let

- A_1=coin 1 matches coin 2
- A_2=coin 2 matches coin 3
- A_3=coin 3 matches coin 1

Let's see if these events are mutually independent. Clearly, $P(A_i) = 2/4 = 1/2$ for all i. Similarly, $P(A_i \cap A_j) = 2/8 = 1/4$ for any i and j, if $i \neq j$, as such an event occurs iff all three coins are head or all of them are tail. Thus $P(A_i) \cdot P(A_j) = P(A_i \cap A_j)$. However, these events are NOT mutually independent, as $P(A_1 \cap A_2 \cap A_3) = 1/4 \neq P(A_1)P(A_2)P(A_3) = 1/8$. This illustrates how important it is to check all subsets when looking for mutual independence.

A weaker condition on independence is the following:

Definition 1.5 *Events $A_1, A_2, \cdots, A_n$ are pairwise independent, if for all i and j, $i \neq j$, A_i and A_j are independent.*

Thus mutual independence implies pairwise independence, but the converse is not true, as the previous example shows.

2 The Birthday Problem

Say there are m=80 people in the classroom. What is the probablility of the event A that there are two of them who have the same birthday?

There are N=365 days in a year. Therefore, we can describe the data in a vector of length m, say $(b_1, b_2, \cdots, b_m)$ so that b_i denotes the birthday of the i-th person. Then we have N choices for each coordinate. The set of these vectors will be our *sample space*.

Definition 2.1 *A sample space S is said to be uniform (or equiprobable) if Pr(S) is the same for all $s \in S$.*

Theorem 2.2 *If S is uniform and $A \subseteq S$, then $Pr(A) = |A|/|S|$.*

Proof. $Pr(A) = \sum_{a \in A} Pr(a) = \sum_{a \in A} 1/|S| = |A|/|S|$ ∎

In the sequel we assume that our sample space is uniform, that is, the birthdays of the people in the classroom are independent from each other. This means that for any $s \in S$, we have $Pr(s) = 1/N^m$ by the previous theorem.

Definition 2.3 *The complement of the event A is S-A. It will be denoted by $\bar{A}$.*

Theorem 2.4

$$Pr(\bar{A}) = 1 - Pr(A)$$

Proof. Obvious. ∎

Now we are computing the number of sample points (vectors) with all different values to get the probability of the event $\bar{A}$ This is easy: we have N choices for the first person's birthday, N-1 for the second person's birthday, and so on, finally we have N-m+1 choices for the last person's birthday. Thus

$$Pr(\bar{A}) = \frac{N(N-1)\cdots(N-m+1)}{N^m}$$

This answers our question, but we are interested to know how large this number is. Sums are usually easier to deal with than products, thus we apply the well known inequality $1-x < e^{-x}$ with $x = i/N$. This yields:

$$Pr(\bar{A}) = \Pi_{i=1}^{m-1}(1 - \frac{i}{N}) \leq \Pi_{i=1}^{m-1} e^{-i/N} =$$

$$= e^{-\sum_i^{m-1} -\frac{i}{N}} = e^{-\frac{(m-1)m}{2N}}$$

To get a lower bound, use the fact that $1 - x > e^{-x-2x^2/3}$ for $0 \leq x \leq 1/3$. This yields

$$Pr(\bar{A}) \geq \Pi_{i=1}^{m-1} e^{-i/N - 2i^2/3N^2} = e^{-\sum_{i=1}^{m-1} i/N + 2i^2/3N^2} \geq$$

$$\geq e^{-\frac{m(m-1)}{2N} + \frac{2m^3}{9N^2}}$$

If we use our upper bound in the case of N=365, m=80, then it follows that there is only 0.02 percent chance not to have two people in the classroom who have the same birthday. This is very surprising, as 80 is not even one fourth of 365.