The Janson Inequality

Geoff Atkinson (following a lecture of Joel Spencer)

December 8, 1998

1 Why do we need the Janson Inequality?

Consider a collection of independent events, $\{A_i\}$. Let $\{B_j\}$ be 'bad' events which we hope to avoid, each the union of one or more of the A_is. For example, if A_i is the event that a certain edge occurs in a given graph, then a 'bad' event, B_j might be a certain arrangement of edges which we are trying to show does not exist in that graph. If all of the B_j events were independent, then it would be easy to compute the probability that no 'bad' events happen using the formula:

$$Pr(\text{No } B_j \text{ occurs}) = \prod(1 - Pr(B_j))$$

If the 'bad' events were only slightly dependent upon one another, then this formula might still give a reasonable approximation. The Janson Inequality gives us a way of formalizing 'only slightly dependent' and measuring just how close the approximation actually is.

2 An Example

Let G be a random graph on n vertices, such that the edge between any two vertices occurs with probability p. Given p, we would like to compute the probability that G is triangle free. In this case, a 'bad' event would be the occurance of a triangle on some specific triplet of vertices. These events are not independent because the existence of a triangle on vertices a, b, and c means that there is more likely to be a triangle on vertices b, c, and d, since the edge from b to c is definitely in the graph.

If the occurances of triangles were independent, then we could compute the probability that G is triangle free as:

$$Pr(G \text{ is triangle free}) = (1 - p^3)^{\binom{n}{3}}$$

The Janson Inequality tells us how closely this approximates the true probability that G is triangle free.

3 A Statement of the Janson Inequality

Let A_i be independent events and let B_j be unions of the A_is.

Let $M = \prod(1 - Pr(B_j))$. This is the probability of no B_j events occuring if all the B_j events were independent.

Let $\Delta = \sum Pr(B_i \wedge B_j)$ for B_i and B_j dependent, and $i \neq j$. If Δ is small then we have a concrete measure of what we mean by saying the B_j events are 'only slightly dependent'.

Choose $\varepsilon \geq Pr(B_j)$ for all j.

Then, $M \leq Pr(\text{No } B_j \text{ occurs}) \leq Me^{\Delta/(1-\varepsilon)2}$

4 Example Revisited

Let us now apply the Janson Inequality to the example for different values of p. Intuitively, we might expect that for small values of p, the triangles will be 'only slightly dependent', while for large values of p the Janson inequality will not be so helpful.

First, consider $p = 1/n$. We can conservatively choose $\varepsilon = 1/2$. $\Delta = 1/n$ so applying the Janson inequality we get:

$$(1-(1/n)^3)^{\binom{n}{3}} \leq Pr(\text{No } B_j \text{ occurs}) \leq (1-(1/n)^3)^{\binom{n}{3}} e^{1/n}$$

In this case, the Janson Inequality gives a tight bound on the probability.

Now, consider $p = 1/\sqrt{n}$. Again, $\varepsilon = 1/2$ is a conservative choice. In this case $\Delta = n^{1.5}$, so applying the Janson Inequality yields:

$$(1-(1/\sqrt{n})^3)^{\binom{n}{3}} \leq Pr(\text{No } B_j \text{ occurs}) \leq (1-(1/\sqrt{n})^3)^{\binom{n}{3}} e^{n^{1.5}}$$

In this case, the Janson Inequality is not helpful in bounding the probability, since $e^{n^{1.5}} \to \infty$ as n becomes large. In general, Janson's Inequality is only truly useful when $\Delta = o(1)$.

5 A Second Example using the Janson Inequality

Once again consider a random graph, G, on n vertices, where each edge has p probability of occuring in the graph. This time we would like to consider the likelihood of there being a length 3 path between any pair of vertices in the graph. Specifically, we would like to prove the claim:

> If $p = (\frac{c \ln n}{n^2})^{1/3}$ and $c > 2$ then with probability approaching 1, there exists a path of length 3 between any two vertices in the graph.

We will prove this claim by first considering V_{xy} the event that there exists no path of length 3 between vertices x and y. We will show, using the Janson Inequality, that $Pr(V_{xy}) = o(n^{-2})$ and, since there are only $\binom{n}{2}$ potential pairs of vertices, x and y, then the probability that any such pair is not connected by a path of length 3 is $o(1)$.

If we are trying to bound V_{xy} using the Janson inequality, then each bad event, B_j, is the existance of a path of length 3 from x to y using two particular intermediate vertices. However, some of these are dependent because two different paths might have an edge in common.

We can compute:
$M = (1-p^3)^{(n-2)(n-3)}$
$\leq e^{-n^2}p^3$
$= e^{-n^2 \frac{c \ln n}{n^2}}$
$= e^{-c \ln n}$
$= n^{-c}$
$= o(n^{-2})$

We also need to compute Δ in order to apply the Janson Inequality. There are three ways in which a pair of length three paths, x-a-b-y and x-c-d-y, can be dependent; either $a = c$, $b = d$, or $a = d$ and $b = c$. Analyzing each of these cases, we find that $\Delta = \binom{n}{3} p^5 + \binom{n}{3} p^5 + \binom{n}{2} p^5$. Since $\Delta = o(1)$ then we can apply Janson's inequality and get a tight bound, $Pr(V_{xy}) = o(n^{-2})$.