Lecture 9

Threshold Phenomena

March 7, 2005
Lecturer: Nati Linial
Notes: Chris Ré

9.1 Asides

There is a good survey of this area by Gil Kalai, Muli Safra called *Threshold Phenomena and Influence* due out very soon.

9.1.1 Percolation

Though our main technical result concerns random graphs in the G(n,p) model, let us mention other contexts in which threshold phenomena occur. One classical example is *Percolation*, an area started in physics. A typical question here is this: given a planar grid and $0 < p < 1$. Create a graph by keeping each edge of the planar grid with probability p and removing each edge with probability 1-p. The inclusion of edges is done independently. Our question is then: In the resulting graph is the origin in an infinite connected component?

It turns out that there is a critical probability, p_c, such that

$p < p_c$	with probability 1, the origin is not in an infinite component
$p > p_c$	with probability > 0, the origin is in an infinite component

You can imagine considering other similiar questions on higher dimensional grids. For the planar grids it turns out that $p_c = \frac{1}{2}$.

This problem comes up in mining in the following idealized model. Somewhere underground is a deposit of oil. It is surrounded by rocks whose structure is that of a 'random sponge', a solid with randomly placed cavities. The question is how far the oil is likely to flow away from its original location. Percolation in a 3-dimensional setting is a good abstraction of the above physical situation.

Now imagine graphing the probability of the property holding versus the p value from above. As an example see figure 9.1. The interesting questions are how does it behave around or slightly to the right of p_c. For example is this a smooth function? Is it differentiable? How large is its derivative? Figure 9.1 illustrates some curves that could happen. In this example, the property could be discontinuous at p_c or is continuous but not smooth at p_c.

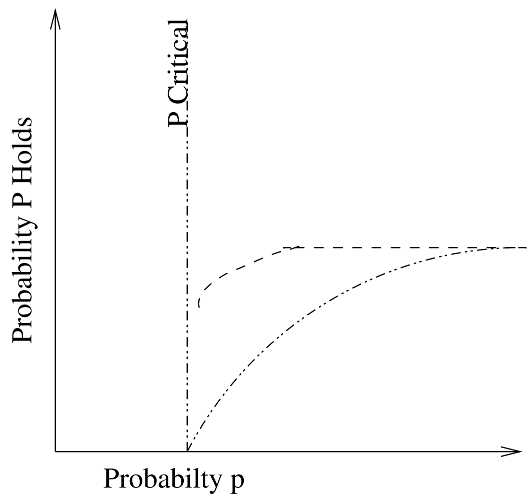

Figure 9.1: Probability of Property vs. p

9.2 Monotone Graph Properties

The main theorem we want to prove is:

Theorem 9.1 (Friedgut and Kalai). *Every monotone graph property has a sharp threshold*

To make this precise, we need some definitions. Let P be a graph property, that is a property invariant under vertex relabeling. A property P is *monotone* if $P(G_0)$ implies that $P(G)$ for all G such that G_0 is a subgraph of G. A property has a sharp threshold, if $Pr[A|G(n,p_1)] = \epsilon$, $Pr[A|G(n,p_2)] = 1-\epsilon$ and $p_2 - p_1 = o(1)$

Theorem 9.2 (Erdös and Renyi). *The threshold for graph connectivity is at* $p = \frac{\log n}{n}$

$p < (1-\epsilon)\frac{\log n}{n}$	G almost surely disconnected
$p > (1+\epsilon)\frac{\log n}{n}$	G almost surely connected

There is a 'counter-point' model to our deleting model, where we throw in edges. There are some surprising facts in this model. For example, when you throw in the edge that reaches the last isolated vertex, with almost certainty, you also connect the graph - at the exact same stage. At the same instant, you also make the graph hamiltonian.

It may be illustrative to see the form of these arguments.

Proof. sketch Let $p < (1-\epsilon)\frac{\log n}{n}$. Let X be a random variable representing the number of isolated vertices. Then $E[X] \to \infty$ since $E[x] = n(1-p)^{n-1}$. We also need a second moment argument like Chebyshev to deduce $X > 0$ almost surely. In particular, when $X > 0$, the graph is disconnected. □

Proof. Let Y_k be a random variable that counts the number of sets $S \subset V$ with $|S| = k$ that have no edges between S and its complement. Then the $E[Y_k] = \binom{n}{k}(1-p)^{k(n-k)}$. It can be checked that if $p > (1+\epsilon)\frac{\log n}{n}$, then $\sum_{k\le\frac{n}{2}} E[Y_k] = o(1)$. It follows that with probability $1-o(1)$ no such sets exist. Clearly, when no such sets exist, the graph is connected. $\square$

9.2.1 Relation to KKL

Why should we expect KKL to work like these examples?

If $f:\{0,1\}^n \to \{0,1\}$ with $E[f]=\frac{1}{2}$[1]. By KKL, $\exists x\in[n]\ Inf_f(x) > \Omega(\frac{\log n}{n})$. Let $N=\binom{[n]}{2}$ then each $z\in\{0,1\}^N$ is a description of an n vertex graph and the variables correspond to edges.

We can now view graph property as an N-variable boolean function. Notice also by symmetry if one edge (variable) is influential, then all edges (variables) are influential. As we will see later large influence entails a sharp threshold.

To generalize, we need to understand the role of p in $G(n,p)$. We have to work with $\{0,1\}^N$ not under the uniform distribution but under the following product distribution: $Pr[U]=p^{|U|}(1-p)^{N-|U|} = p^{E(G)}(1-p)^{\binom{n}{2}-E(G)}$. We are denoting the Hamming weight of U as $|U|$. and $E(G)$ is the edge set of the graph G.

9.3 BK^3L

9.3.1 A relation between influence and the derivative of $\mu_p(A)$

The new B and K in our theory are Bourgain and Katznelson. By $\mu_p(A)$ we denote the probability that the property A holds under the $p, 1-p$ product measure.

Lemma 9.3 (Margulis & Russo). *Let $A\subseteq\{0,1\}^n$ be a monotone subset and let $\mu_p(A)$ be the p-measure of A. For $x\in A$ let $h(x)=|\{y\notin A|x,y\in E(cube)\}|$ (number of neigbors of x outside of A).*

Let $\Psi_p(A)=\sum h(x)\mu_p(x)$, the weighted sum of these hs.

Additionally let $\Phi_p(A)$ be the sum of influences of individual variables. Then

$$\Phi_p(A) =_1 \frac{\Psi_p(A)}{p} =_2 \frac{d}{dp}\mu_p(A)$$

The subscripts on the equality are only for convenience in the proof.

Definition 9.1. We will say $x\succ y$, if x and y differ in exactly one coordinate, say the i^{th}, and $x_i=1$ and $y_i=0$.

Influences, more generally In general, if X is a probability space and if $f: X^n\to\{0,1\}$ (i.e. f can be viewed as an indicator function for a subset of X^n). For $1\le k\le n$, we can say $Inf_f(k) = Pr_{X^{n-1}}[$ Obtain a non-constant fiber $]$. Here we are randomly choosing $n-1$ coordinates from X with the k^{th} coordinate missing, and checking if the resulting fiber is constant for f. Namely, if the value of f is fixed regardless of the choice of for the k^{th} variable.

[1] choosing $E[f]=\frac{1}{2}$ is not critical. Anything bounded away from 0,1 will do

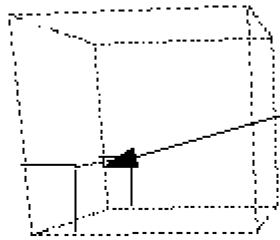

Figure 9.2: Cube with a Fiber

Proof. We prove equality 1. $\Phi_p(A)$ is the sum of all influences. The influence of the i^{th} variable is the weighted sum of all such edges such that $x \succ y$ where $x \in A$, $y \notin A$ and $x_i = 1, y_i = 0$. The probability of the relevent event is this: We have selected all coordinates except the i^{th} and the outcome should coincide with x. There are $|x| - 1$ coordinates which are 1 among those and $n - |x|$ coordinates for which are 0. So we can rewrite the formula as follows

$$\begin{aligned}
\Phi_p(A) &= \sum_{x \in A, y \notin A, x \succ y} p^{(|x|-1)}(1-p)^{n-|x|} \\
&= \frac{1}{p} \sum_{x \in A, y \notin A, x \succ y} p^{(|x|)}(1-p)^{n-|x|} \\
&= \frac{1}{p} \sum_{x \in A} p^{(|x|)}(1-p)^{n-|x|} \, |\{y | y \notin A, x \succ y\}| \\
&= \frac{1}{p} \sum_{x \in A} p^{(|x|)}(1-p)^{n-|x|} \, h(x) \\
&= \frac{1}{p} \sum_{x \in A} \mu_p(x) h(x) = \frac{1}{p} \Psi_p(A)
\end{aligned}$$

□

Proof. Equality 2.

$$\frac{d}{dp}\mu_p(A) = \sum_{x \in A} |x| p^{|x|-1}(1-p)^{n-|x|} - \sum_{x \in A} (n - |x|) p^{|x|}(1-p)^{n-|x|-1}$$

$$p\frac{d}{dp}\mu_p(A) = \sum_{x \in A} |x| p^{|x|}(1-p)^{n-|x|} - \frac{p}{1-p} \sum_{x \in A} (n - |x|) p^{|x|}(1-p)^{n-|x|}$$

For a fixed vertex of the cube, x, and e an edge incident with x define

$$w_{x,e} = \begin{cases} 1 & \text{e goes down from x} \\ -\frac{p}{1-p} & \text{e goes up from x} \end{cases}$$

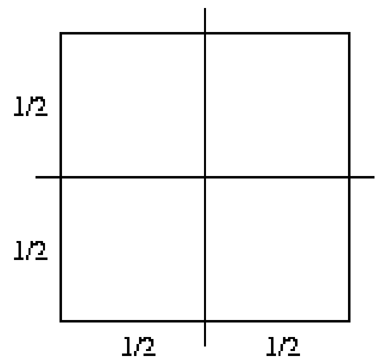

Figure 9.3: Partitioning the cube to derive KKL from BK^3L

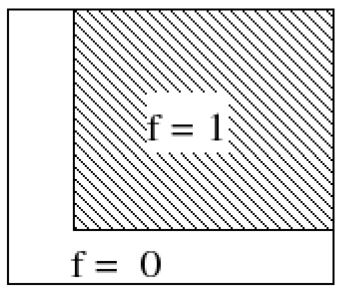

Figure 9.4: f on n=2 case

So we can rewrite (summing over x and e's incident).

$$p\frac{d}{dp}\mu_p(A) = \sum_{x\in A, e\sim x} w_{x,e}\mu_p(x)$$

This is because there are $|x|$ edges going down from x and $|n-x|$ edges going up from it. Notice that if $x \succ y$ are both in A and $e=(x,y)$, then $w_{x,e}\mu_p(x) + w_{y,e}\mu_p(y) = 0$. It follows that we can restrict to the sum to the edges in $E(A, A^c)$. In other words,

$$p\frac{d}{dp}\mu_p(A) = \sum_{x\in A, y\notin A, e=(x,y)} w_{x,e}\ \mu_p(x) = \sum_{x\in A, y\notin A, e=(x,y)} \mu_p(x) = \sum_{x\in A} h(x)\mu_p(x) = \Psi_p(A)$$

□

Returning to the proof that every monotone graph property has a sharp threshold. Let A be a monotone graph property and let us operatre in the probability space $G(n,p)$. We will show here that the p value where the property holds with less than ϵ is very close to where the property holds with $\frac{1}{2}$. A symmetric argument for $1-\epsilon$ will give us the full desired result.

9.3.2 Words about BKKKL

Theorem 9.4 (BKKKL). *Let $f : [0,1]^n \to \{0,1\}$ with $E[f] = t$, let $t' = min(t, 1-t)$. Then there exists $n \geq k \geq 1$ such that $Inf_f(k) \geq \Omega(t' \frac{\log n}{n})$*

Set version of KKL $\forall f : \{0,1\}^n \to \{0,1\}$ $E[f] \sim \frac{1}{2}$ and for every $\omega(n) \to \infty$ as $n \to \infty$, $\exists S \subseteq [n]$. $|S| \leq \frac{n(\omega(n))}{\log n}$ with $Inf_f(S) = 1 - o(1)$. This result follows from repeated application of KKL.

Remark. It is interesting to note that the analgous statement for $f : [0,1]^n \to [0,1]$ does not hold.

Consider the following f, represented in figure 9.4. Let $f(x_1, \ldots, x_n) = 0$ iff $\exists i$ $0 \leq x_i \leq \frac{c}{n}$ where $c = \log_e(2)$. In other words, $f^{-1}(1) = \prod_i^n [\frac{c}{n}, 1]$. Let $|S| = \alpha$. In this example, $Inf(S) = Pr[$f still undetermined when all variables outside of S are set at random$]$. The function is still undetermined iff all others outside the set are 1. This happens with probability $(1 - \frac{c}{n})^{n(1-\alpha)} \approx e^{-c(1-\alpha)}$, which is bounded away from 1.

This is a 'close-cousin' of the tribes example. Recall in the tribes example we broke the variables into 'tribes' of size $\sim \log n - \log\log n$. Each tribe contributed if all variables take on the value 1, that is there is one assignment out of the $2^{\log n - \log\log n} = \frac{n}{\log n}$ such that the tribe had value 1. In our setting, we can identify tribes with single variables. The 0 region of the continuous case corresponds to the assigment where all variables in the discrete case are set to 1, since this determines the function.

Proof. By BK^3L there exist influential variables. By symmetry all variables are influential. Sum of all individual influences are at least as large as $\Phi_p(A) \geq \Omega(\epsilon \ \log N) = \Omega(\epsilon \ \log n)$

$$\Phi_p(A) \geq \Omega(\mu_p(A) \ \log n)$$

By Margulis-Russo Lemma we know $\Phi_p(A) = \frac{d}{dp} \ \mu_p(A)$.

$$\frac{d}{dp}\mu_p(A) \geq \Omega(\mu_p(A) \ \log n)$$

$$(\frac{d}{dp}\mu_p(A))/\mu_p(A) \geq \Omega(\log n)$$

$$\frac{d}{dp}(\log(\mu_p(A)) \geq \Omega(\log n)$$

let p_1, p_2 be defined by $Pr_{G(n,p_1)}[A] = \epsilon$ and $Pr_{G(n,p_2)}[A] = \frac{1}{2}$. From above we know that $d(\log \mu_p(A)) > \Omega(\log n)$ so $p_1 - p_2 < O(\frac{\log \frac{1}{\epsilon}}{\log n})$. $\square$

Remark. We will not give a proof here, but note that Freidgut showed using standard measure theory how to derive BK^3L from KKL. Namely, how we can reach the same conclusion for any $f : X^n \to \{0,1\}$, where X is any probability space. To derive KKL from BK^3L, is easy: Given $f : \{0,1\}^n \to \{0,1\}$ define $F : [0,1]^n \to \{0,1\}$ by breaking the cube to 2^n subcubes and letting F be constant on each subcube that is equal to f at the correpsoning vertex of the cube. For a simple illustration of the case $n = 2$, see figure 9.3.