Amitabha Bagchi
Scribe: Arpit Nanda

Lecture 3: Increasing events and the FKG inequality

20th August 2007

We now move on to one of the important tools that will be used repeatedly in the study of percolation: the FKG inequality. Before we do, we identify a particular class of events which will come up again and again as we proceed.

3.1 Increasing events

Recall the partial order defined on the sample space Ω in a previous lecture: For $\omega_1, \omega_2 \in \Omega$ we say that $\omega_1 \leq \omega_2$ if

$$\forall e \in \mathbb{E}^2, \omega_1(e) \leq \omega_2(e).$$

$$\Rightarrow \text{ Any edge open in } \omega_1, \text{must be open in } \omega_2.$$

Now consider an event $A \subseteq \Omega$ defined with respect to two vertices x and y as follows

A : There is an open path from x to y.

Notice that for some $\omega_1 \in A$ and $\omega_2 \in \Omega$, if $\omega_1 \leq \omega_2$ then $\omega_2 \in A$, since the edges which are open in ω_1, $K(w_1)$ are enough to guarantee the desired path, and the edges open in ω_2, are a superset of $K(w_1)$.

Now consider another event $B \subseteq \Omega$ defined on two verices u and v:

B : There is a closed path from u to v.

Now here if $\omega_1 \in B$ then some ω_2 satisfying $\omega_1 \leq \omega_2$ may or may not be in B. But here we are able to claim that if $\omega_2 \in B$ then $\forall \omega_1 \leq \omega_2 : \omega_1 \in B$.

Now consider an event $C \subseteq \Omega$ defined on vertices w and z

C : There is path with alternating open and closed edges from w to z.

For this event C it is the case that if $\omega_1 \in C$ there may be an ω_2 such that $\omega_1 \leq \omega_2$ and $\omega_2 \notin C$. And there may be $\omega_3 < \omega_1$ such that $\omega_3 \notin C$.

Motivated by this discussion we give the following definition:

Definition 3.1 *$A \in \mathcal{F}$ is called an* increasing event *if $I_A(\omega) \leq I_A(\omega')$, whenever $\omega \leq \omega'$, where I_A is the indicator variable of A.*

Example 3.2 *Some examples of increasing events:*

- *The origin is contained in an infinite cluster.*
- *There is an open path from x to y.*
- *The edges of $B(n)$ are open.*

Similarly, we have that a random variable N is called an *increasing* random variable if for $\omega_1, \omega_2 \in \Omega$, $N(\omega_1) \leq N(\omega_2)$ whenever $\omega_1 \leq \omega_2$.

Example 3.3 *Some examples of increasing random variables:*

- *The size of the open cluster containing vertex x.*
- *The largest k such that all the edges in $B(k)$ are open.*

Intuitively it appears that the probability that an increasing event occurs should increase as the probability of edges remaining open increases. This is in fact true.

Theorem 3.4 *For an increasing event A, and an increasing random variable N, given probabilities $p_1 \leq p_2$,*

$$P_{p_1}(A) \leq P_{p_2}(A), \tag{1}$$
$$E_{p_1}(N) \leq E_{p_2}(N). \tag{2}$$

Proof. We use a coupling argument to prove this theorem, noting that this coupling argument will be used often as we proceed with the study of percolation. We will only prove the first statement, leaving the second as an exercise.

For each $e \in \mathbb{E}^d$, let us define a random variables $X(e)$ distributed uniformly at random in $[0, 1]$. Furthermore, for a probability value p, we define a quantity $\eta_p(e)$ as follows

$$\eta_p(e) = \begin{cases} 1 & \text{if } X(e) \leq p \\ 0 & \text{otherwise} \end{cases}$$

What we are doing here is using the variables $X(e)$ as the basic random experiment from which we can generate percolation processes for whatever value of p we want. The value of $\eta_p(e)$ tells us whether the edge e is open or closed in the percolation process with probability p generated this way.

Now, let us use this method to generate two percolation processes, one for p_1 and the other for p_2. Note that the random experiment of choosing the values of the $X(e)$s is conducted only once. Now consider the outcomes generated for these two processes:

$$\omega_{p_1} = \eta_{p_1}(e_1)\eta_{p_1}(e_2)\ldots$$

$$\omega_{p_2} = \eta_{p_2}(e_1)\eta_{p_2}(e_2)\ldots$$

So, it is easy to see that $\omega_{p_1} \leq \omega_{p_2}$ if $p_1 \leq p_2$. And since A is an increasing event $\omega_{p_1} \in A$ implies that $\omega_{p_2} \in A$ which proves the result. ■.

In an earlier lecture we proved using Kolmogorov's 0-1 law that the percolation probability $\theta(p)$ is either 0 or 1. Now, with Theorem 3.4 we are in a position to say that the graph of the percolation probability does not look like Figure 1. It looks like Figure 2.

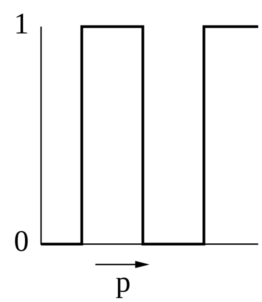

Figure 1: An incorrect picture of the percolation probability.

3.2 The FKG Inequality

Given the definition of increasing events, it is natural to think that two increasing events might be positively correlated. The FKG inequality shows precisely this.

Theorem 3.5 FKG inequality. *If A and B are increasing events then*

$$P_p(A \cap B) \geq P_p(A) \cdot P_p(B).$$

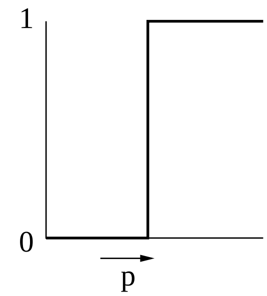

Figure 2: A correct picture of the percolation probability.

Also, if X and Y are increasing random variables

$$E_p(XY) \geq E_p(X) \cdot E_p(Y).$$

Proof. We will not give the whole proof here, just a flavour of it.

Let X and Y be two increasing r.v. Assume that they only depend on one edge e,which has two states 0 and 1. Take two values γ_1 and γ_2 each of which can have 2 values: 1 and 1. Now, since X and Y are increasing r. v. the signs of $X(\gamma_1) - X(\gamma_2)$ and $Y(\gamma_1) - X(\gamma_2)$ are the same no matter what values γ_1 and γ_2 take. Hence

$$[X(\gamma_1) - X(\gamma_2)][Y(\gamma_1) - Y(\gamma_2)] \geq 0$$

Summing over all values of γ_1 and γ_2 we get

$$\sum_{\gamma_1,\gamma_2 \in \{1,0\}} [X(\gamma_1) - X(\gamma_2)][Y(\gamma_1) - Y(\gamma_2)] \mathrm{P}_p(\omega(e) = \gamma_1) \mathrm{P}_p(\omega(e) = \gamma_2) > 0.$$

Some manipulation shows that this is equivalent to saying that

$$2(\mathrm{E}_p(XY) - \mathrm{E}_p(X)\mathrm{E}_p(Y)) \geq 0.$$

Which proves the result for increasing random variables which depend on only one edge. The rest of the proof proceeds by using this as the base case of an induction which extends this to increasing random variables that depend only on a finite number of edges. Some results from analysis help us extend it to all increasing random variables. We omit those details, referring the reader to [2].

To prove the first statement of the theorem we consider the special random variables, I_A and I_B which are the indicator random variables of the events A and B. Since $\mathrm{E}_p(I_A) = \mathrm{P}_p(A)$, the result follows. ■

Example 3.6 *The FKG inequality can be used to prove that the probability that a vertex lies in an infinite cluster is translation invariant.*

Consider an infinite mesh with points x and y. Let $\theta(p, x)$ be the probability that x lies in an infinite cluster. Then $p_c(x) = sup\{p : \theta(p, x) = 0\}$. Also, let the ad-hoc notation $x \leftrightarrow \infty$ mean that x is a part of an infinite cluster.

Now, if x is part of an infinite cluster, it could be that y is also part of that cluster, or it could be that y is not part of that cluster. In general we can say that

$$x \leftrightarrow \infty \supseteq (x \leftrightarrow y) \cap (y \leftrightarrow \infty).$$

This means that

$$\mathrm{P}_p(x \leftrightarrow \infty) \geq \mathrm{P}_p[(x \leftrightarrow y) \cap (y \leftrightarrow \infty)]$$

Now, since both $x \leftrightarrow y$ and $y \leftrightarrow \infty$ are increasing events we can use the FKG inequality and get that

$$\mathrm{P}_p(x \leftrightarrow \infty) \geq \mathrm{P}_p(x \leftrightarrow y) \cdot \mathrm{P}_p(y \leftrightarrow \infty).$$

Hence if $\mathrm{P}_p(x \leftrightarrow \infty)$ is 0 then $\mathrm{P}_p(y \leftrightarrow \infty)$ must be 0 since $\mathrm{P}_p(x \leftrightarrow y)$ is non-zero. And this means that

$$p_c(y) \geq p_c(x).$$

Exactly the same argument can be made with x and y interchanged, thereby proving the following theorem.

Theorem 3.7 *For any two vertices x and y of $\mathbb{L}^d$*

$$p_c(x) = p_c(y).$$

3.3 The general form of the FKG inequality

The FKG inequality was actually proved for a more general setting than what has been described above. In this section we give the general form of the FKG inequality and an application of it to a simple occupancy problem. The reader only interested in percolation can safely skip this section.

We start by introducing some notation. A *lattice* is a partially ordered set in which each pair of elements, x and y, has a unique minimal upper bound, called the *join* of x and y, denoted $x \vee y$, and a unique maximal lower bound, called the *meet* of x and y and denoted $x \wedge y$. A lattice L is said to be *distributive* if, for all $x, y, z \in L$

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$

Given a distributive lattice L, a function $\mu : L \to \mathbb{R}_+$ is called *log-supermodular* if

$$\mu(x) \cdot \mu(y) \leq \mu(x \vee y) \cdot \mu(x \wedge y)$$

for all $x, y \in L$.

A function $f : L \to \mathbb{R}$ is *non-decreasing* if $f(x) \leq f(y)$ whenever $x \leq y$ and *non-increasing* if $f(x) \geq f(y)$ whenever $x \leq y$.

In this setting, we have the following statement of the FKG inequality:

Theorem 3.8 (FKG Inequality) *Let L be a finite distributive lattice and let $\mu : L \to \mathbb{R}_+$ be a log-supermodular funtion. Then if $f, g : L \to \mathbb{R}_+$ are both non-decreasing or non-increasing, we have*

$$\left(\sum_{x \in L} \mu(x) f(x)\right) \cdot \left(\sum_{x \in L} \mu(x) g(x)\right) \leq \left(\sum_{x \in L} \mu(x) f(x) g(x)\right) \cdot \left(\sum_{x \in L} \mu(x)\right)$$

If we take μ to be a measure on L. Assuming μ to not be identically 0, we can define the expectation w.r.t. μ of a function f as

$$\mathrm{E}_\mu[f] = \frac{\sum_{x \in L} \mu(x) f(x)}{\sum_{x \in L} \mu(x)}.$$

With this notation, we can restate the FKG inequality as follows: For any log-supermodular funtion μ defined on a distributive lattice and any two functions $f : L \to \mathbb{R}$ and $g : L \to \mathbb{R}$ which are both non-increasing or non-decreasing,

$$\mathrm{E}_\mu[f \cdot g] \geq \mathrm{E}_\mu[f] \cdot \mathrm{E}_\mu[g].$$

And, if one of them is non-increasing and the other non-decreasing,

$$\mathrm{E}_\mu[f \cdot g] \leq \mathrm{E}_\mu[f] \cdot \mathrm{E}_\mu[g].$$

3.4 A simple application of the general FKG inequality

The application described in this section is taken from the work of Dubhashi and Ranjan [1] as discussed in [3].

Suppose we throw m balls independently into n bins uniformly at random, for positive integers m and n. Let X_i be the random variable denoting the number of balls in the ith bin. We want to show the intuitively obvious result that X_i and X_j are negatively correlated. And we want to do this without any calculation.

Let us begin by proving the following lemma

Lemma 3.9 *Given two bins, i and j, and two positive integers $t_i, t_j \leq m$,*

$$P((X_i \geq t_i) \wedge (X_j \geq t_j)) \leq P(X_i \geq t_i) \cdot P(X_j \geq t_j).$$

Proof. We will use the general version of the FKG inequality for this.

In order to create an appropriate lattice, we use an m-dimensional vector $\hat{a} = (a_1, a_2, \ldots, a_m)$ to denote each outcome of the experiment, where a_i is the bin number (between 1 and n) of the bin into which the ith ball landed. Let L be the set of all such outcomes.

Define a partial order: $\hat{a} \leq_L \hat{b}$ if $a_i \leq b_i$ for all $i \in [m]$.

Suppose $x = (x_1, ..., x_m)$ and $y = (y_1, ..., y_m)$ are two m-dimensional vectors in L. Define $u = (u_1, ..., u_m)$ such that $u_i = max(x_i, y_i)$ for $1 \leq i \leq m$. It is easy to verify that $u = x \vee y$. Similarly, if we define $v = (v_1, ..., v_m)$ such that $v_i = min(x_i, y_i)$ for $1 \leq i \leq m$, it is easy to verify that $v = x \wedge y$.

Clearly, u and v are unique for a given pair x and y. Hence, L is a lattice.

To show that L is a distributive lattice, we will prove the following simple fact:

Fact 3.10 *If a, b, c are three real numbers, then*

$$min(a, max(b, c)) = max(min(a, b), min(a, c)).$$

Proof. There are two cases: $b \leq c$ or $b > c$. Since the cases are symmetric, we will assume that $b \leq c$. The other case is similar. In this case, $max(b, c) = c$.

Hence, the LHS $= min(a,c)$. For $b \leq c, min(a,b) \leq min(a,c)$. Hence, the RHS $= min(a,c)$. ■

From the above fact it follows that

$$\min(x_i, \max(y_i, z_i)) = \max(\min(x_i, y_i), \min(x_i, z_i)) \text{ for } 1 \leq i \leq m.$$

Note that the LHS is the i-th component of the m-dimensional vector $x \wedge (y \vee z)$, whereas the RHS is the i-th component of the m-dimensional vector $(x \wedge y) \vee (x \wedge z)$. It follows that,

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

Hence, L is a distributive lattice.

It is easy to see that μ is trivially log-supermodular since $\mu(x) \cdot \mu(y) = 1/n^m \cdot 1/n^m = 1/n^{2m}$ and similarly $\mu(x \vee y) \cdot \mu(x \wedge y) = 1/n^m \cdot 1/n^m = 1/n^{2m}$.

Given any two bins i and j, note that we can renumber the bins so that bin i becomes bin 1 and bin j becomes bin n.

Consider now the events $A = \{X_1 \geq t_i\}$ and $B = \{X_n \geq t_j\}$. Given the way the lattice is (partially) ordered, it is clear that A is a decreasing event and B is an increasing event. This is because if $x \leq y$ for some $x, y \in L$, this means that for each ball i $x_i \leq y_i$ i.e. if there were some balls in bin 1 in x, in y they could have moved out to some other bin with a higher number. However, if we consider the balls which were already in bin n, they cannot move to a higher bin number. On the other hand, balls with a smaller bin number can move into bin n.

Hence, using FKG inequality, the result follows. ■

We leave it to the reader to show that Lemma 3.9 can be used to show that X_i and X_j are negatively correlated.

References

[1] D. Dubhashi and D. Ranjan. Some correlation inequalities for probabilistic analysis of algorithms. Technical Report MPI-I-94-143, Max Planck Institut für Informatik, Saarbrücken, Germany, 1994.

[2] G. Grimmett. *Percolation.* Springer, 1999.

[3] C. Scheideler. Introduction to probability theory. Manuscript.