7 Correlation Inequalities

7.1 Harris–FKG inequality

Recall that $A \subseteq \{0,1\}^n$ is called an ***increasing event*** (also: ***increasing property***, ***up-set***) if A is upwards-closed, meaning that whenever x is in A, then everything above x in the boolean lattice also lies in A. In other words,

$$\text{if } x \in A \text{ and } x \leq y \text{ (coordinatewise), then } y \in A.$$

Similarly, a ***decreasing event*** is defined by a downward closed collection of subset of $\{0,1\}^n$. A subset $A \subseteq \{0,1\}^n$ is increasing if and only if its complement $\overline{A} \subseteq \{0,1\}^n$ is decreasing.

The main theorem of this chapter tells us that

increasing events of independent variables are positively correlated .

Theorem 7.1.1 (Harris 1960)
If A and B are increasing events of independent boolean random variables, then

$$\mathbb{P}(AB) \geq \mathbb{P}(A)\mathbb{P}(B).$$

Equivalently, we can write $\mathbb{P}(A \mid B) \geq \mathbb{P}(A)$.

***Remark* 7.1.2** (Independence assumption)**.** It is important that the boolean random variables are independent, also they do not have to be identically distributed.

There are other important settings where the independence assumption can be relaxed. This is important for certain statistical physics models, where much of this theory originally arose. Indeed, the above inequality is often called the ***FKG inequality***, attributed to Fortuin, Kasteleyn, Ginibre (1971), who proved a more general result in the setting of distributive lattices, which we will not discuss here (see Alon–Spencer).

***Remark* 7.1.3** (Percolation)**.** Many of such inequalities were initially introduced for the study of *percolations*. A classic setting of this problem takes place in infinite grid with vertices $\mathbb{Z}^2$ with edges connecting adjacent vertices at distance 1. Suppose we keep each edge of this infinite grid with probability p independently, what is the probability that the origin is part of an infinite component (in which case we say that there is "percolation")? This is supposed to an idealized mathematical model of how a fluid permeates through a medium. Harris showed that with probability 1, percolation does not occur for $p \leq 1/2$. A later breakthrough of Kesten (1980) shows that percolation occurs with probability 1 for all $p > 1/2$. Thus the "bond percolation threshold" for $\mathbb{Z}^2$ is exactly $1/2$. Such exact results are extremely rare.

***Example* 7.1.4.** Here is a quick application of Harris' inequality to a random graph $G(n, p)$:

$$\mathbb{P}(\text{planar} \mid \text{connected}) \leq \mathbb{P}(\text{planar}).$$

Indeed, being planar is a decreasing property, whereas being connected is an increasing property.

We state and prove a more general result, which says that independent random variables possess ***positive association***.

Let each Ω_i be a linearly ordered set (i.e., $\{0, 1\}$, $\mathbb{R}$) and $x_i \in \Omega_i$ with respect to some probability distribution independent for each i. We say that a function $f(x_1, \dots, x_n)$ is ***monotone increasing*** if

$$f(x) \leq f(y) \text{ whenever } x \leq y \text{ coordinatewise.}$$

Theorem 7.1.5 (Harris)
If f and g are monotone increasing functions of independent random variables, then

$$\mathbb{E}[fg] \geq (\mathbb{E}f)(\mathbb{E}g).$$

This version of Harris inequality implies the earlier version by setting $f = 1_A$ and $g = 1_B$.

Proof. We use induction on n.

For $n = 1$, for independent $x, y \in \Omega_1$, we have

$$0 \leq \mathbb{E}[(f(x) - f(y))(g(x) - g(y))] = 2\mathbb{E}[fg] - 2(\mathbb{E}f)(\mathbb{E}g).$$

So $\mathbb{E}[fg] \geq (\mathbb{E}f)(\mathbb{E}g)$. (The one-variable case is sometimes called Chebyshev's inequality. It can also be deduced using the rearrangement inequality).

Now assume $n \geq 2$. Let $h = fg \colon \Omega_1 \times \cdots \times \Omega_n \to \mathbb{R}$. Define marginals $f_1, g_1, h_1 \colon \Omega_1 \to$

$\mathbb{R}$ by

$$\begin{aligned}
f_1(y_1) &= \mathbb{E}[f|x_1 = y_1] = \mathbb{E}_{(x_2,\dots,x_n)\in\Omega_2\times\cdots\times\Omega_n}[f(y_1,x_2,\dots,x_n)],\\
g_1(y_1) &= \mathbb{E}[g|x_1 = y_1] = \mathbb{E}_{(x_2,\dots,x_n)\in\Omega_2\times\cdots\times\Omega_n}[g(y_1,x_2,\dots,x_n)],\\
h_1(y_1) &= \mathbb{E}[h|x_1 = y_1] = \mathbb{E}_{(x_2,\dots,x_n)\in\Omega_2\times\cdots\times\Omega_n}[h(y_1,x_2,\dots,x_n)].
\end{aligned}$$

Note that f_1 and g_1 are 1-variable monotone increasing functions on Ω_1.

For every fixed $y_1 \in \Omega_1$, the function $(x_2,\dots,x_n) \mapsto f(y_1,x_2,\dots,x_n)$ is monotone increasing, and likewise with g. So applying the induction hypothesis for $n-1$, we have

$$h_1(y_1) \geq f_1(y_1)g_1(y_1). \tag{7.1}$$

Thus

$$\begin{aligned}
\mathbb{E}[fg] = \mathbb{E}[h] = \mathbb{E}[h_1] &\geq \mathbb{E}[f_1g_1] && \text{[by (7.1)]}\\
&\geq (\mathbb{E}f_1)(\mathbb{E}g_1) && \text{[by the } n=1 \text{ case]}\\
&= (\mathbb{E}f)(\mathbb{E}g).
\end{aligned}$$

□

Corollary 7.1.6 (Decreasing events and multiple events)
Let A and B be events on independent random variables.

(a) If A and B are decreasing, then $\mathbb{P}(A \wedge B) \geq \mathbb{P}(A)\mathbb{P}(B)$.

(b) If A is increasing and B is decreasing, then $\mathbb{P}(A \wedge B) \leq \mathbb{P}(A)\mathbb{P}(B)$.

If $A_1,\dots,A_k$ are all increasing (or all decreasing) events on independent random variables, then

$$\mathbb{P}(A_1\cdots A_k) \geq \mathbb{P}(A_1)\cdots\mathbb{P}(A_k).$$

Proof. For the second inequality, note that the complement $\overline{B}$ is increasing, so

$$\mathbb{P}(AB) = \mathbb{P}(A) - \mathbb{P}(A\overline{B}) \overset{\text{Harris}}{\leq} \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(\overline{B}) = \mathbb{P}(A)\mathbb{P}(B).$$

The proof of the first inequality is similar. For the last inequality we apply the Harris inequality repeatedly. □

7.2 Applications to random graphs

Triangle-free probability

Question 7.2.1

What's the probability that $G(n, p)$ is triangle-free?

Harris inequality will allow us to prove a lower bound. In the next chapter, we will use Janson inequalities to derive upper bounds.

Theorem 7.2.2

$\mathbb{P}(G(n, p) \text{ is triangle-free}) \geq (1 - p^3)^{\binom{n}{3}}$

Proof. For each triple of distinct vertices $i, j, k \in [n]$, the event that ijk does not form a triangle is a decreasing event (here the ground set is the set of edges of the complete graph on n). So by Harris' inequality,

$$\begin{aligned}\mathbb{P}(G(n, p) \text{ is triangle-free}) &= \mathbb{P}\left(\bigwedge_{i<j<k} \{ijk \text{ not a triangle}\}\right) \\ &\geq \prod_{i<j<k} \mathbb{P}(ijk \text{ not a triangle}) = (1 - p^3)^{\binom{n}{3}}. \qquad \square\end{aligned}$$

***Remark* 7.2.3.** How good is this bound? For $p \leq 0.99$, we have $1 - p^3 = e^{-\Theta(p^3)}$, so the above bound gives

$$\mathbb{P}(G(n, p) \text{ is triangle-free}) \geq e^{-\Theta(n^3 p^3)}.$$

Here is another lower bound

$$\mathbb{P}(G(n, p) \text{ is triangle-free}) \geq \mathbb{P}(G(n, p) \text{ is empty}) = (1 - p)^{\binom{n}{2}} = e^{-\Theta(n^2 p)}.$$

The bound from Harris is better when $p \ll n^{-1/2}$. Putting them together, we obtain

$$\mathbb{P}(G(n, p) \text{ is triangle-free}) \gtrsim \begin{cases} e^{-\Theta(n^3 p^3)} & \text{if } p \lesssim n^{-1/2} \\ e^{-\Theta(n^2 p)} & \text{if } n^{-1/2} \lesssim p \leq 0.99 \end{cases}$$

(note that the asymptotics agree at the boundary $p \asymp n^{-1/2}$). In the next chapter, we will prove matching upper bounds using Janson inequalities.

Maximum degree

Question 7.2.4

What's the probability that the maximum degree of $G(n, 1/2)$ is at most $n/2$?

For each vertex v, $\deg(v) \le n/2$ is a decreasing event with probability just slightly over $1/2$. So by Harris inequality, the probability that every v has $\deg(v) \le n/2$ is at least $\ge 2^{-n}$.

It turns out that the appearance of high degree vertices is much more correlated than the independent case. The truth is exponentially more than the above bound.

Theorem 7.2.5 (Riordan and Selby 2000)

$$\mathbb{P}(\operatorname{maxdeg} G(n, 1/2) \le n/2) = (0.6102 \cdots + o(1))^n$$

Instead of giving a proof, we consider an easier continuous model of the problem that motivates the numerical answer. Building on this intuition, Riordan and Selby (2000) proved the result in the random graph setting, although this is beyond the scope of this class.

In a random graphs, we assign independent Bernoulli random variables on edges of a complete graph. Instead, let us assign independent standard normal random variables to each edge of the complete graph.

Proposition 7.2.6 (Max degree with normal random edge labels)

Assign an independent standard normal random variable Z_{uv} to each edge of K_n. Let $W_v = \sum_{u \ne v} Z_{uv}$ be the sum of the labels of the edges incident to a vertex v. Then

$$\mathbb{P}(W_v \le 0 \; \forall v) = (0.6102 \cdots + o(1))^n$$

The event $W_v \le 0$ is supposed to model the event that the degree at vertex v is less than $n/2$. Of course, other than intuition, there is no justification here that these two models should behave similarly

We have $\mathbb{P}(W_v \le 0) = 1/2$. Since each $\{W_v \le 0\}$ is a decreasing event of the independent edge labels, Harris' inequality tells us that

$$\mathbb{P}(W_v \le 0 \; \forall v) \ge 2^{-n}.$$

The truth turns out to be significantly greater.

Proof sketch of Proposition 7.2.6. The tuple $(W_v)_{v \in [n]}$ has a joint normal distribution, with each coordinate variance $n - 1$ and pairwise covariance 1. So $(W_v)_{v \in [n]}$ has the same distribution as

$$\sqrt{n-2}(Z_1', Z_2', \ldots, Z_n') + Z_0'(1, 1, \ldots, 1)$$

where $Z_0', \ldots, Z_n'$ are iid standard normals.

Let Φ be the pdf and cdf of the standard normal $N(0,1)$.

Thus

$$\mathbb{P}(W_v \le 0 \,\forall v) = \mathbb{P}\left(Z_i' \le -\frac{Z_0'}{\sqrt{n-2}} \;\forall i \in [n]\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} \Phi\left(\frac{-z}{\sqrt{n-2}}\right)^n dz$$

where the final step is obtained by conditioning on Z_0'. Substituting $z = y\sqrt{n}$, the above quantity equals to

$$= \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\infty} e^{nf(y)}\, dy \quad \text{where} \quad f(y) = -\frac{y^2}{2} + \log \Phi\left(y\sqrt{\frac{n}{n-2}}\right).$$

We can estimate the above integral for large n using the *Laplace method* (which can be justified rigorously by considering Taylor expansion around the maximum of f). We have

$$f(y) \approx g(y) := -\frac{y^2}{2} + \log \Phi(y)$$

and we can deduce that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\max_{v \in [n]} W_v \le 0) = \lim_{n\to\infty} \frac{1}{n} \log \int e^{nf(y)}\, dy = \max g = \log 0.6102 \cdots. \qquad \square$$

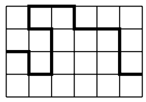
Exercises

1. Let $G = (V, E)$ be a graph. Color every edge with red or blue independently and uniformly at random. Let E_0 be the set of red edges and E_1 the set of blue edges. Let $G_i = (V, E_i)$ for each $i = 0, 1$. Prove that

$$\mathbb{P}(G_0 \text{ and } G_1 \text{ are both connected}) \le \mathbb{P}(G_0 \text{ is connected})^2.$$

2. A set family $\mathcal{F}$ is *intersecting* if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{F}$. Let $\mathcal{F}_1, \ldots, \mathcal{F}_k$ each be a collection of subsets of $[n]$ and suppose that each $\mathcal{F}_i$ is intersecting. Prove that $\left|\bigcup_{i=1}^k \mathcal{F}_i\right| \le 2^n - 2^{n-k}$.

3. *Percolation.* Let $G_{m,n}$ be the grid graph on vertex set $[m] \times [n]$ (m vertices wide and n vertices tall). A *horizontal crossing* is a path that connects some left-most vertex to some right-most vertex. See below for an example of a horizontal crossing in $G_{7,5}$.



Let $H_{m,n}$ denote the random subgraph of $G_{m,n}$ obtained by keeping every edge with probability $1/2$ independently.

Let $\mathsf{RSW}(k)$ denote the following statement: there exists a constant $c_k > 0$ such that for all positive integers n, $\mathbb{P}(H_{kn,n}$ has a horizontal crossing$) \geq c_k$.

a) Prove $\mathsf{RSW}(1)$.

b) Prove that $\mathsf{RSW}(2)$ implies $\mathsf{RSW}(100)$.

c) $\star\star$ (Very challenging) Prove $\mathsf{RSW}(2)$.

4. Let A and B be two *independent* increasing events of independent random variables. Prove that there are two *disjoint* subsets S and T of these random variables so that A depends only on S and B depends only on T.

5. Let U_1 and U_2 be increasing events and D a decreasing event of independent Boolean random variables. Suppose U_1 and U_2 are independent. Prove that $\mathbb{P}(U_1|U_2 \cap D) \leq \mathbb{P}(U_1|U_2)$.

6. *Coupon collector.* Let $s_1, \ldots, s_m$ be independent random elements in $[n]$ (not necessarily uniform or identically distributed; chosen with replacement) and $S = \{s_1, \ldots, s_m\}$. Let I and J be disjoint subsets of $[n]$. Prove that $\mathbb{P}(I \cup J \subseteq S) \leq \mathbb{P}(I \subseteq S)\mathbb{P}(J \subseteq S)$.

7. $\star$ Prove that there exist $c < 1$ and $\varepsilon > 0$ such that if $A_1, \ldots, A_k$ are increasing events of independent Boolean random variables with $\mathbb{P}(A_i) \leq \varepsilon$ for all i, then, letting X denote the number of events A_i that occur, one has $\mathbb{P}(X = 1) \leq c$. (Give your smallest c. It is conjectured that any $c > 1/e$ works.)

8. $\star$ *Disjoint containment.* Let $\mathcal{S}$ and $\mathcal{T}$ each be a collection of subsets of $[n]$. Let $R \subseteq [n]$ be a random subset where each element is included independently (not necessarily with the same probability). Let A be the event that $S \subseteq R$ for some $S \in \mathcal{S}$. Let B be the event that $T \subseteq R$ for some $T \in \mathcal{T}$. Let C denote the event there exist *disjoint* $S, T \subseteq R$ with $S \in \mathcal{S}$ and $T \in \mathcal{T}$. Prove that $\mathbb{P}(C) \leq \mathbb{P}(A)\mathbb{P}(B)$.