7. Correlation Inequalities

- Let $n \in \mathbb{N}$ and consider a probability space $(\Omega, \Pr)$ whose sample points (elementary events) correspond to subsets of $[n] = \{1, \dots, n\}$, i.e. $\Omega = \mathcal{P}([n])$.
  - E.g. for m-vertex random graphs, $n = \binom{m}{2}$, and each $A \subseteq [n]$ represents a specific realization of the edges in the graph.
- An event $\mathcal{A} \subseteq \Omega$ is <u>monotone</u> <u>increasing</u> if $A \in \mathcal{A}$, $A \subseteq A' \Rightarrow A' \in \mathcal{A}$. (Resp. <u>monotone</u> <u>decreasing</u> if $A \in \mathcal{A}$, $A' \subseteq A \Rightarrow A' \in \mathcal{A}$.)
  - E.g. many interesting graph properties are monotone increasing: existence of specific subgraphs, connectedness, Hamiltonicity, ...
- Intuitively, two monotone increasing (decreasing) events $\mathcal{A}, \mathcal{B}$ should "support" each other, so that

$$\Pr(\mathcal{A} \mid \mathcal{B}) \geq \Pr(\mathcal{A}) \quad , \quad \text{i.e. } \Pr(\mathcal{A} \cap \mathcal{B}) \geq \Pr(\mathcal{A}) \cdot \Pr(\mathcal{B}).$$

Possibly also an increasing event $\mathcal{A}$ and a decreasing event $\mathcal{C}$ should "oppose" each other, so that

$$\Pr(\mathcal{A} \mid \mathcal{C}) \leq \Pr(\mathcal{A}), \quad \text{i.e. } \Pr(\mathcal{A} \cap \mathcal{C}) \leq \Pr(\mathcal{A}) \cdot \Pr(\mathcal{C}).$$

- Very general (and eventually simple) results of this kind proved by Kleitman (1966), Ahlswede & Daykin (1978) and Fortuin, Kasteleyn & Ginibre (1971).

7.1 The Ahlswede - Daykin "Four Functions Theorem"

- Let $n \in \mathbb{N}$ and consider any "weight" function $\varphi : \mathcal{P}([n]) \to \mathbb{R}^+$ on the subsets of $[n]$. For $\mathcal{A} \subseteq \mathcal{P}([n])$, denote

$$\varphi(\mathcal{A}) = \sum_{A \in \mathcal{A}} \varphi(A).$$

For $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$, denote

$$\mathcal{A} \vee \mathcal{B} = \{ A \cup B \mid A \in \mathcal{A},\ B \in \mathcal{B} \},$$

$$\mathcal{A} \wedge \mathcal{B} = \{ A \cap B \mid A \in \mathcal{A},\ B \in \mathcal{B} \}.$$

- Theorem 7.1 (A & D 1978). Let $\alpha, \beta, \gamma, \delta : \mathcal{P}([n]) \to \mathbb{R}^+$ be any four weight functions on the subsets of $[n]$. If for any two subsets $A, B \subseteq [n]$:

(*) $\alpha(A)\beta(B) \le \gamma(A \cup B)\delta(A \cap B)$,

then for any two families of subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$:

(**) $\alpha(\mathcal{A})\beta(\mathcal{B}) \le \gamma(\mathcal{A} \vee \mathcal{B})\delta(\mathcal{A} \wedge \mathcal{B})$.

Proof. Assume that (*) holds and consider any two set families $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$. We may in fact assume that $\mathcal{A} = \mathcal{B} = \mathcal{A} \vee \mathcal{B} = \mathcal{A} \wedge \mathcal{B} = \mathcal{P}([n])$. For if not, modify the weight functions so that

$$\alpha(A) = 0 \text{ for } A \notin \mathcal{A}, \qquad \beta(B) = 0 \text{ for } B \notin \mathcal{B}$$
$$\gamma(C) = 0 \text{ for } C \notin \mathcal{A} \vee \mathcal{B}, \quad \delta(D) = 0 \text{ for } D \notin \mathcal{A} \wedge \mathcal{B}.$$

Then (*) still holds for the modified $\alpha, \beta, \gamma, \delta$, and (**) holds w.r.t. $\mathcal{P}([n])$ iff it holds w.r.t. the original families $\mathcal{A}, \mathcal{B}$.

- Let us then consider the case $n=1$. Then $\mathcal{P}([n]) = \{\emptyset, \{1\}\}$. For each $\varphi \in \{\alpha, \beta, \gamma, \delta\}$, denote $\varphi_0 = \varphi(\emptyset)$, $\varphi_1 = \varphi(\{1\})$. By $(*)$, we have

$$(+) \quad \alpha_0\beta_0 \leq \gamma_0\delta_0, \quad \alpha_0\beta_1 \leq \gamma_1\delta_0, \\ \alpha_1\beta_0 \leq \gamma_1\delta_0, \quad \alpha_1\beta_1 \leq \gamma_1\delta_1.$$

By the simplification, we only need to consider $(**)$ in the case $\mathcal{A} = \mathcal{B} = \{\emptyset, \{1\}\}$, i.e. to prove

$$(++) \quad (\alpha_0+\alpha_1)(\beta_0+\beta_1) \leq (\gamma_0+\gamma_1)(\delta_0+\delta_1).$$

- If $\gamma_1 = 0$ or $\delta_0 = 0$, $(++)$ follows directly from $(+)$.

Otherwise, by $(+)$, $\gamma_0 \geq \frac{\alpha_0\beta_0}{\delta_0}$ and $\delta_1 \geq \frac{\alpha_1\beta_1}{\gamma_1}$.

To establish $(++)$ in this case, it thus suffices to show:

$$\left(\frac{\alpha_0\beta_0}{\delta_0} + \gamma_1\right)\left(\delta_0 + \frac{\alpha_1\beta_1}{\gamma_1}\right) \geq (\alpha_0+\alpha_1)(\beta_0+\beta_1)$$

$$\Leftrightarrow (\alpha_0\beta_0 + \gamma_1\delta_0)(\delta_0\gamma_1 + \alpha_1\beta_1) \geq (\alpha_0+\alpha_1)(\beta_0+\beta_1)\delta_0\gamma_1$$

$$\Leftrightarrow (\gamma_1\delta_0 - \alpha_0\beta_1)(\gamma_1\delta_0 - \alpha_1\beta_0) \geq 0.$$

But the last inequality is ostensibly true, because both factors on the left-hand side are nonnegative.

This completes the proof for $n=1$.

- For the case $n \geq 2$, assume by induction that the result holds for $n-1$. Now for each $\varphi : \mathcal{P}([n]) \to \mathbb{R}^+$, $\varphi \in \{\alpha, \beta, \gamma, \delta\}$, define an associated $\varphi' : \mathcal{P}([n-1]) \to \mathbb{R}^+$ by

$$\varphi'(A') = \varphi(A') + \varphi(A' \cup \{n\}).$$

Clearly for $\mathcal{A} = \mathcal{P}([n])$, $\mathcal{A}' = \mathcal{P}([n-1])$: $\varphi(\mathcal{A}) = \varphi'(\mathcal{A}')$.

Thus, assuming that $\alpha, \beta, \gamma, \delta : \mathcal{P}([n]) \to \mathbb{R}^+$ satisfy condition $(+)$, we only need to show that also the associated $\alpha', \beta', \gamma', \delta' : \mathcal{P}([n-1]) \to \mathbb{R}^+$ satisfy condition $(+)$ and apply the induction hypothesis.

To validate $(+)$ for the functions $\alpha', \beta', \gamma', \delta'$ and a given pair of sets $A', B' \subseteq [n-1]$ consider again the powerset of some 1-element set $\{t\}$ and define:

$$\begin{aligned}
\bar{\alpha}(\emptyset) &= \alpha(A'), & \bar{\alpha}(\{t\}) &= \alpha(A' \cup \{n\}) \\
\bar{\beta}(\emptyset) &= \beta(B'), & \bar{\beta}(\{t\}) &= \beta(B' \cup \{n\}) \\
\bar{\gamma}(\emptyset) &= \gamma(A' \cup B'), & \bar{\gamma}(\{t\}) &= \gamma(A' \cup B' \cup \{n\}) \\
\bar{\delta}(\emptyset) &= \delta(A' \cap B'), & \bar{\delta}(\{t\}) &= \delta((A' \cap B') \cup \{n\}).
\end{aligned}$$

By the assumption that $\alpha, \beta, \gamma, \delta$ satisfy $(+)$, also:

$$\bar{\alpha}(S)\bar{\beta}(T) \leq \bar{\gamma}(S \cup T)\,\bar{\delta}(S \cap T) \qquad \text{for } S, T \subseteq \{t\}.$$

Hence, by re-applying the result for the case $n = 1$:

$$\alpha'(A')\beta'(B') = (\alpha(A') + \alpha(A' \cup \{n\}))(\beta(B') + \beta(B' \cup \{n\}))$$

$$= \bar{\alpha}(\{\emptyset, \{t\}\})\,\bar{\beta}(\{\emptyset, \{t\}\})$$

$$\leq \bar{\gamma}(\{\emptyset, \{t\}\})\,\bar{\delta}(\{\emptyset, \{t\}\})$$

$$= (\gamma(A' \cup B') + \gamma(A' \cup B' \cup \{n\}))\,(\delta(A' \cap B') + \delta((A' \cap B') \cup \{n\}))$$

$$= \gamma'(A' \cup B')\,\delta'(A' \cap B').$$

This completes the proof. $\square$

• Corollary 7.2 For any two families of sets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}([n])$:

$$|\mathcal{A}| \cdot |\mathcal{B}| \leq |\mathcal{A} \vee \mathcal{B}| \cdot |\mathcal{A} \wedge \mathcal{B}|.$$

Proof. Choose $\alpha = \beta = \gamma = \delta \equiv 1$ in Thm 7.1. $\square$

• Corollary 7.3 For any set family $\mathcal{F} \subseteq \mathcal{P}([n])$, denote

$$\mathcal{F} \setminus \mathcal{F} = \{ A \setminus B \mid A \in \mathcal{F}, B \in \mathcal{F} \}.$$

Then $|\mathcal{F} \setminus \mathcal{F}| \geq |\mathcal{F}|$.

Proof. Apply Cor. 7.2 to $\mathcal{F}$ and $\mathcal{F}' = \{ [n] \setminus A \mid A \in \mathcal{F} \}$:

$$|\mathcal{F}|^2 = |\mathcal{F}| \cdot |\mathcal{F}'| \leq |\mathcal{F} \vee \mathcal{F}'| \cdot |\mathcal{F} \wedge \mathcal{F}'| = |\mathcal{F} \setminus \mathcal{F}|^2. \quad \square$$

7.2 The FKG Theorem

• Let $n \in \mathbb{N}$ and denote for brevity $\Omega = \mathcal{P}([n])$.

• A function ("measure") $\mu : \Omega \to \mathbb{R}^+$ is log-supermodular if if for all $A, B \subseteq [n]$:

$$\mu(A)\mu(B) \leq \mu(A \cup B)\mu(A \cap B).$$

A function $f : \Omega \to \mathbb{R}^+$ is increasing if $A \subseteq B \Rightarrow f(x) \leq f(y)$ and decreasing if $A \subseteq B \Rightarrow f(x) \geq f(y)$.

- Theorem 7.4 (Fortuin, Kasteleyn & Ginibre 1971).
Let $\mu: \Omega \to \mathbb{R}^+$ be a log-supermodular measure and $f, g: \Omega \to \mathbb{R}^+$ two increasing functions. Then:

$$\Big(\sum_A \mu(A)f(A)\Big)\cdot\Big(\sum_A \mu(A)g(A)\Big) \leq \Big(\sum_A \mu(A)f(A)g(A)\Big)\cdot\Big(\sum_A \mu(A)\Big).$$

Proof. Define four functions $\alpha, \beta, \gamma, \delta$ as follows:

$$\alpha(A) = \mu(A)f(A) \qquad \beta(A) = \mu(A)g(A)$$
$$\gamma(A) = \mu(A)f(A)g(A) \qquad \delta(A) = \mu(A).$$

Then, by the supermodularity of μ and since f, g are increasing:

$$\begin{aligned}\alpha(A)\beta(B) &= \mu(A)f(A)\mu(B)g(B)\\ &\leq \mu(A\cup B)f(A)g(B)\mu(A\cap B)\\ &\leq \mu(A\cup B)f(A\cup B)g(A\cup B)\mu(A\cap B)\\ &= \gamma(A\cup B)\,\delta(A\cap B).\end{aligned}$$

By the Ahlswede-Daykin Theorem (Thm 7.1), thus:

$$\alpha(\Omega)\beta(\Omega) \leq \gamma(\Omega)\,\delta(\Omega),$$

which is the desired result. $\square$

- Thm 7.4 holds also if both f and g are decreasing (interchange γ and δ in the proof). If f is increasing and g decreasing (or vice versa), then the opposite inequality holds:

$$\Big(\sum_A \mu(A)f(A)\Big)\cdot\Big(\sum_A \mu(A)g(A)\Big) \geq \Big(\sum_A \mu(A)f(A)g(A)\Big)\cdot\Big(\sum_A \mu(A)\Big).$$

(Apply Thm 7.4 to the two increasing functions $f(x)$ and $k - g(x)$, where $k = \max_{A\in[n]} g(A)$.)

- Assuming that the measure $\mu : \Omega \to \mathbb{R}^+$ is not $\equiv 0$, one may define an "expectation" w.r.t. it:

$$E_\mu[f] = \frac{\sum_A \mu(A) f(A)}{\sum_A \mu(A)}.$$

Then Thm 7.4 states that if μ is log-supermodular and f, g are both increasing or decreasing, then:

$$E_\mu[fg] \geq E_\mu[f] \cdot E_\mu[g]$$

and if f is increasing and g decreasing (or vice versa), then:

$$E_\mu[fg] \leq E_\mu[f] \cdot E_\mu[g].$$

- Let us then get back to our original application of pairs of monotone increasing (decreasing) events.

- Let $n \in \mathbb{N}$ and associate to each $i \in [n]$ an "elementary choice probability" p_i, $0 \leq p_i \leq 1$. Define a measure μ on $\Omega = \mathcal{P}([n])$ by:

$$\mu(A) = \prod_{i \in A} p_i \cdot \prod_{j \notin A} (1 - p_j).$$

It is easy to check that μ is in fact a probability measure on Ω and moreover log-supermodular. (In fact, for any $A, B \subseteq [n]$: $\mu(A)\mu(B) = \mu(A \cup B)\mu(A \cap B)$ as can be seen by comparing the contributions of each $i \in [n]$ to the lhs and rhs of the equation.)

- Now for a monotone increasing (decreasing) event, the characteristic function χ is increasing (decreasing). Thus, applying the FKG theorem to the char. functions yields:

indicator

- Corollary 7.5 Let μ be the prob measure on $\Omega = \mathcal{P}([n])$ defined above and $\mathcal{A}, \mathcal{B}$ increasing and $\mathcal{C}, \mathcal{D}$ decreasing events. Then: $\mu(\mathcal{A} \cap \mathcal{B}) \geq \mu(\mathcal{A}) \cdot \mu(\mathcal{B})$, $\mu(\mathcal{C} \cap \mathcal{D}) \geq \mu(\mathcal{C}) \cdot \mu(\mathcal{D})$, $\mu(\mathcal{A} \cap \mathcal{C}) \leq \mu(\mathcal{A}) \cdot \mu(\mathcal{C})$. $\square$