Algebraic methods toward higher-order probability inequalities

Donald Richards

Penn State University and SAMSI

Background

L: A finite distributive lattice with partial order $\preccurlyeq$

Least upper bound: $\vee$

Greatest lower bound: $\wedge$

L is order-isomorphic to the lattice of subsets of a finite set

There exists a finite set A such that $L = 2^A$, the set of all subsets of A

2^A is a finite distributive lattice
Partial order: $\subseteq$
Least upper bound: $\cup$
Greatest lower bound: $\cap$

μ: A probability measure on 2^A

μ is *multivariate totally positive of order* 2 (MTP$_2$) if

$$\mu(a \cup b)\mu(a \cap b) \geq \mu(a)\mu(b)$$

for all $a, b \subseteq A$

Other terminology:

log-supermodular

FKG measure

A function $f : 2^A \rightarrow \mathbb{R}$ is *increasing* if

$$f(a) \leq f(b), \qquad a \subseteq b$$

The *expected value* of f with respect to μ:

$$\mathbb{E}(f) := \sum_{a \subseteq A} \mu(a) f(a)$$

The FKG inequality (Fortuin, Kasteleyn, Ginibre, 1971):

If $f_1, f_2 : L \rightarrow \mathbb{R}$ are increasing andμ is MTP$_2$ then

$$\text{Cov}(f_1, f_2) := \mathbb{E}(f_1 f_2) - \mathbb{E}(f_1)\mathbb{E}(f_2) \geq 0$$

Applications

Statistics: Multivariate statistical analysis, dependence properties of random variables, observational studies

Probability theory: Diffusion equations, reliability theory, percolation

Mathematical physics: Interacting particle systems, Ising models

Total positivity, finite reflection groups, analysis on Lie groups

Combinatorics: Monotonicity of partial orders, Sperner theory, graph theory, Ramsey theory

The case of $\mathbb{R}^n$

$x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$

The partial order: $x \preccurlyeq y$ if $x_j \leq y_j$, $j = 1, \ldots, n$

The lattice operations

$$x \vee y = (\max(x_1, y_1), \ldots, \max(x_n, y_n))$$
$$x \wedge y = (\min(x_1, y_1), \ldots, \min(x_n, y_n))$$

$f : \mathbb{R}^n \to \mathbb{R}$ is *increasing* if $f(x) \leq f(y)$ whenever $x \preccurlyeq y$

A probability density function $K : \mathbb{R}^n \longrightarrow \mathbb{R}$ is MTP$_2$ if

$$K(x \vee y)K(x \wedge y) \geq K(x)K(y), \qquad x, y \in \mathbb{R}^n$$

A remarkable result: For suitably smooth K, the MTP$_2$ condition is equivalent to

$$\frac{\partial^2}{\partial x_i \partial x_j} \ln K(x) \geq 0 \quad \text{for all } i \neq j$$

The FKG inequality: If $f_1, f_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ are increasing and K is MTP$_2$ then

$$\int_{\mathbb{R}^n} f_1 f_2 K \, dx - \Big(\int_{\mathbb{R}^n} f_1 K \, dx \Big) \cdot \Big(\int_{\mathbb{R}^n} f_2 K \, dx \Big) \geq 0$$

The normal distribution

$(X_1, \ldots, X_n) \sim \mathcal{N}_n(0, \Sigma)$

$$K(x) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\big(-\tfrac{1}{2} x' \Sigma^{-1} x \big)$$

K is MTP$_2$ if $(\Sigma^{-1})_{ij} \leq 0$ for all $i \neq j$

For $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, construct the indicator function

$$I_a(x) = \begin{cases} 1, & x_1 \geq a_1, \ldots, x_n \geq a_n \\ 0, & \text{otherwise} \end{cases}$$

I_a is increasing

If $(X_1, \ldots, X_n)$ is a MTP$_2$ normal random vector then

$$\begin{aligned}&\mathbb{E} I_a(X_1, \ldots, X_n) I_b(X_1, \ldots, X_n)\\ &\qquad \geq \mathbb{E} I_a(X_1, \ldots, X_n) \cdot \mathbb{E} I_b(X_1, \ldots, X_n)\end{aligned}$$

Equivalently,

$$\begin{aligned}&P(X_1 \geq a_1 \vee b_1, \ldots, X_n \geq a_n \vee b_n)\\ &\qquad \geq P(X_1 \geq a_1, \ldots, X_n \geq a_n) \cdot P(X_1 \geq b_1, \ldots, X_n \geq b_n)\end{aligned}$$

There are *many* applications of this result in statistical inference

Cumulants

$X \in \mathbb{R}^n$: A random vector

$f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$

Cumulants: The coefficients in the Taylor-Maclaurin expansion of the *cumulant-generating function*

$$\log \mathbb{E} \exp(t_1 f_1 + \cdots + t_n f_m)$$

$\kappa_2 = \mathbb{E}(f_1 f_2) - \mathbb{E}(f_1)\mathbb{E}(f_2)$ is the simplest cumulant

The next three cumulants are

$$\begin{aligned}\kappa_3 := {} & \mathbb{E}(f_1 f_2 f_3) - [\mathbb{E}(f_1 f_2)\mathbb{E}(f_3) + \mathbb{E}(f_1 f_3)\mathbb{E}(f_2) + \mathbb{E}(f_1)\mathbb{E}(f_2 f_3)] \\ & + 2\mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)\end{aligned}$$

$$\begin{aligned}\kappa_4 := {} & \mathbb{E}(f_1 f_2 f_3 f_4) - [\mathbb{E}(f_1 f_2 f_3)\mathbb{E}(f_4) + \cdots] - [\mathbb{E}(f_1 f_2)\mathbb{E}(f_3 f_4) + \cdots] \\ & + 2[\mathbb{E}(f_1 f_2)\mathbb{E}(f_3)\mathbb{E}(f_4) + \cdots] - 6\mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\end{aligned}$$

$$\begin{aligned}\kappa_5 := {} & \mathbb{E}(f_1 f_2 f_3 f_4 f_5) - [\mathbb{E}(f_1 f_2 f_3 f_4)\mathbb{E}(f_5) + \cdots] \\ & - [\mathbb{E}(f_1 f_2 f_3)\mathbb{E}(f_4 f_5) + \cdots] \\ & + 2[\mathbb{E}(f_1 f_2 f_3)\mathbb{E}(f_4)\mathbb{E}(f_5) + \cdots] + 2[\mathbb{E}(f_1 f_2)\mathbb{E}(f_3 f_4)\mathbb{E}(f_5) + \cdots] \\ & - 6[\mathbb{E}(f_1 f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\mathbb{E}(f_5) + \cdots] + 24\mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\end{aligned}$$

Question: Are there FKG inequalities for the cumulants?

Answer: Probably not

For the normal distribution, cumulants of order 3 or more are identically zero

It is not difficult to find MTP$_2$ random vectors with negative cumulants

Percus (1975), Sylvester (1975): Interesting correlation inequalities for cumulants

Evidence of cumulant-type inequalities were obtained with indicator functions

Recall that

$$\kappa_3 := \mathbb{E}(f_1 f_2 f_3) - [\mathbb{E}(f_1 f_2)\mathbb{E}(f_3) + \mathbb{E}(f_1 f_3)\mathbb{E}(f_2) + \mathbb{E}(f_1)\mathbb{E}(f_2 f_3)] \\ + 2\mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)$$

Define the *conjugate cumulant*

$$\kappa_3' := 2\mathbb{E}(f_1 f_2 f_3) - [\mathbb{E}(f_1 f_2)\mathbb{E}(f_3) + \mathbb{E}(f_1 f_3)\mathbb{E}(f_2) + \mathbb{E}(f_1)\mathbb{E}(f_2 f_3)] \\ + \mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)$$

“Conjugate”: Reverse the order of the absolute value of the coefficients in the cumulant

Analogy with the conjugate of a partition

L: finite distributive lattice

μ: An MTP$_2$ probability measure on L

f_1, f_2 and f_3: Nonnegative increasing functions on L

After years of hard work ...

Theorem: $\kappa_3'(f_1, f_2, f_3) \geq 0$

My gratitude to the Institute for Advanced Study

Corollary: The FKG inequality

Proof: Set $f_3 \equiv 1$

The theorem cannot be deduced from the FKG inequality

κ_3' is an alternating sum

$$\kappa_3' = \text{Cov}(f_1 f_2, f_3) - \mathbb{E}(f_1)\text{Cov}(f_2, f_3) + \text{Cov}(f_1 f_3, f_2),$$

κ_4 and κ_5: Conjugating the coefficients still works

$f_1, \ldots, f_5$: Nonnegative increasing functions on L

Define the fourth- and fifth-order conjugate cumulants

$$\begin{aligned}\kappa_4' := {} & 6\mathbb{E}(f_1f_2f_3f_4) - 2[\mathbb{E}(f_1f_2f_3)\mathbb{E}(f_4) + \cdots] - [\mathbb{E}(f_1f_2)\mathbb{E}(f_3f_4) + \cdots] \\ & + [\mathbb{E}(f_1f_2)\mathbb{E}(f_3)\mathbb{E}(f_4) + \cdots] - \mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\end{aligned}$$

$$\begin{aligned}\kappa_5' := {} & 24\mathbb{E}(f_1f_2f_3f_4f_5) - 6[\mathbb{E}(f_1f_2f_3f_4)\mathbb{E}(f_5) + \cdots] \\ & - 2[\mathbb{E}(f_1f_2f_3)\mathbb{E}(f_4f_5) + \cdots] \\ & + 2[\mathbb{E}(f_1f_2f_3)\mathbb{E}(f_4)\mathbb{E}(f_5) + \cdots] + [\mathbb{E}(f_1f_2)\mathbb{E}(f_3f_4)\mathbb{E}(f_5) + \cdots] \\ & - [\mathbb{E}(f_1f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\mathbb{E}(f_5) + \cdots] + \mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)\mathbb{E}(f_4)\end{aligned}$$

Theorem: Under the same hypotheses as before, $\kappa_4', \kappa_5' \geq 0$

The proof that $\kappa_3' \geq 0$

We adapt a proof of the FKG inequality due to den Hollander and Keane (1986)

The proof is by induction on the length of maximal chains in L

$L = 2^A$ for some finite set A

If $A = \emptyset$ then the result is trivial, so assume that $A \neq \emptyset$

WLOG, assume that $\mu(a) > 0$ for all $a \subseteq A$

Why? $\mathbb{E}f$ depends only on the support of μ

Choose and fix B, an arbitrarily chosen subset of A

For $a \subseteq B$, define

$$\mu_B(a) := \sum_{b \subseteq A \setminus B} \mu(a \cup b) \qquad (1)$$

μ_B is the marginal probability measure on the lattice 2^B

Easy fact (well-known to statisticians!):

μ_B is MTP$_2$

For $f : 2^A \to \mathbb{R}$, define

$$f_B(a) := \frac{1}{\mu_B(a)} \sum_{b \subseteq A \setminus B} \mu(a \cup b) f(a \cup b) \tag{2}$$

$f_B(a)$ is the conditional expectation of f, given $a \subseteq B$

Easy fact (well-known to statisticians!):

If f is increasing then so is f_B

For $g : 2^B \to \mathbb{R}$, define

$$\mathbb{E}_B(g) := \sum_{a \subseteq B} \mu_B(a) g(a)$$

The Double-Expectation Theorem (well-known to undergrad. statistics & probability students!):

For $B \subseteq A$,

$$\mathbb{E}(f) = \sum_{a \subseteq A} \mu(a) f(a) = \sum_{a \subseteq B} \mu_B(a) f_B(a) = \mathbb{E}_B(f_B)$$

In words: The overall average value of f equals the average value of its conditional averages

Suppose $B = A\backslash\{z\}$ where $z \in A$ is chosen arbitrarily

Shorthand notation: f_{iB} denotes $(f_i)_B$, $i = 1, 2, 3$

The Claim:

$$\begin{aligned} & 2\mathbb{E}_B\big((f_1f_2f_3)_B\big) \\ & \quad - \Big[\mathbb{E}_B\big((f_1f_2)_B f_{3B}\big) + \mathbb{E}_B\big((f_1f_3)_B f_{2B}\big) + \mathbb{E}_B\big(f_{1B}(f_2f_3)_B\big)\Big] \\ & \qquad + \mathbb{E}_B(f_{1B}f_{2B}f_{3B}) \geq 0 \end{aligned} \tag{3}$$

For $a \subseteq B$, it follows from (1) that

$$\mu_B(a) = \mu(a) + \mu(a \cup \{z\}) \tag{4}$$

and, from (2), we get

$$f_B(a) = \frac{1}{\mu_B(a)} \Big(\mu(a) f(a) + \mu(a \cup \{z\}) f(a \cup \{z\}) \Big) \tag{5}$$

Conclude:

$$\begin{aligned} &\mu_B(a)^3 (f_1 f_2 f_3)_B(a) \\ &\qquad = \mu_B(a)^2 \Big[\mu(a) f_1(a) f_2(a) f_3(a) \\ &\qquad\qquad + \mu(a \cup \{z\}) f_1(a \cup \{z\}) f_2(a \cup \{z\}) f_3(a \cup \{z\}) \Big] \end{aligned} \tag{6}$$

For $\{i,j,k\} = \{1,2,3\}$,

$$\begin{aligned}&\mu_B(a)^3 (f_i f_j)_B(a) f_{kB}(a) \\ &\quad = \mu_B(a)\big[\mu(a) f_i(a) f_j(a) + \mu(a\cup\{z\}) f_i(a\cup\{z\}) f_j(a\cup\{z\})\big] \\ &\qquad\qquad \times \big[\mu(a) f_k(a) + \mu(a\cup\{z\}) f_k(a\cup\{z\})\big] \qquad (7)\end{aligned}$$

$$\begin{aligned}&\mu_B(a)^3 f_{1B}(a) f_{2B}(a) f_{3B}(a) \\ &\qquad = \prod_{i=1}^{3} \big[\mu(a) f_i(a) + \mu(a\cup\{z\}) f_i(a\cup\{z\})\big] \qquad (8)\end{aligned}$$

Collect together (6) - (8), simplify the algebraic expressions using (4) and (5):

$$\begin{aligned}\mu_B(a)^3\Big\{&2(f_1 f_2 f_3)_B(a)\\&-\left[(f_1 f_2)_B(a)\, f_{3B}(a) + (f_1 f_3)_B(a)\, f_{2B}(a) + f_{1B}(a)\,(f_2 f_3)_B(a)\right]\\&+ f_{1B}(a) f_{2B}(a) f_{3B}(a)\Big\}\\=\ &\mu(a)\mu(a\cup\{z\})\\&\times\Big[\big(f_1(a\cup\{z\}) - f_1(a)\big)\Phi_{1B}(a) + f_1(a\cup\{z\})\Phi_{2B}(a)\Big]\end{aligned} \tag{9}$$

where

$$
\begin{aligned}
\Phi_{1B}(a) &= \mu(a)\big(f_2(a\cup\{z\})f_3(a\cup\{z\}) - f_2(a)f_3(a)\big) \\
&\quad +\mu(a\cup\{z\})f_3(a)\big(f_2(a\cup\{z\}) - f_2(a)\big) \\
&\quad +\mu(a\cup\{z\})f_2(a)\big(f_3(a\cup\{z\}) - f_3(a)\big)
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{2B}(a) &= \big(\mu(a\cup\{z\}) + \mu(a)\big) \\
&\quad \times\big(f_2(a\cup\{z\}) - f_2(a)\big)\,\big(f_3(a\cup\{z\}) - f_3(a)\big)
\end{aligned}
$$

Each f_i is nonnegative and increasing

Therefore Φ_{1B}, Φ_{2B} are sums of products of nonnegative terms

Conclude: (9) is nonnegative

Divide both sides of (9) by $\mu_B(a)^2$ and sum over all $a \subseteq B$

Double-Expectation Theorem gives:

$$\sum_{a \subseteq B} \mu_B(a)(f_1 f_2 f_3)_B(a) = \mathbb{E}_B\big((f_1 f_2 f_3)_B\big) = \mathbb{E}(f_1 f_2 f_3),$$

For $\{i, j, k\} = \{1, 2, 3\}$,

$$\sum_{a \subseteq B} \mu_B(a)(f_i f_j)_B(a) f_{kB}(a) = \mathbb{E}_B\big((f_i f_j)_B f_{kB}\big)$$

Also,

$$\sum_{a \subseteq B} \mu_B(a) f_{1B}(a) f_{2B}(a) f_{3B}(a) = \mathbb{E}_B(f_{1B} f_{2B} f_{3B})$$

Collect together these identities, apply the nonnegativity of (9); we obtain The Claim:

$$2\mathbb{E}_B\big((f_1 f_2 f_3)_B\big) - \Big[\mathbb{E}_B\big((f_1 f_2)_B f_{3B}\big) + \mathbb{E}_B\big((f_1 f_3)_B f_{2B}\big) + \mathbb{E}_B\big(f_{1B}(f_2 f_3)_B\big)\Big] + \mathbb{E}_B(f_{1B} f_{2B} f_{3B}) \geq 0$$

Finally, set $B = \emptyset$; observe that

$$\mathbb{E}_\emptyset(f_{i\emptyset}) \equiv \mathbb{E}(f_i);\ \mathbb{E}_\emptyset((f_i f_j)_\emptyset f_{k\emptyset}) \equiv \mathbb{E}(f_i f_j)\mathbb{E}(f_k);$$
$$\mathbb{E}_\emptyset(f_{1\emptyset} f_{2\emptyset} f_{3\emptyset}) \equiv \mathbb{E}(f_1)\mathbb{E}(f_2)\mathbb{E}(f_3)$$

The Claim reduces to: $\kappa_3' \geq 0$ Q.E.D.

The proof of $\kappa'_4, \kappa'_5 \geq 0$: Similar

We used MAPLE to carry out the algebraic computations

What about κ'_6? We had some evidence that $\kappa'_6 \not\geq 0$

Siddhartha Sahi (Rutgers Univ.) has found stronger evidence recently

Nevertheless, there is a conjecture ...

A *partition* $\lambda = (\lambda_1, \lambda_2, \ldots)$ is a sequence of nonnegative integers with $\lambda_1 \geq \lambda_2 \geq \cdots$

The *parts* of λ are the non-zero λ_i

The *weight* of λ is $|\lambda| := \lambda_1 + \lambda_2 + \cdots$

The *length* of λ, denoted by $l(\lambda)$, is the number of parts of λ

For $i = 1, 2, \ldots$, let λ'_i denote the cardinality of the set $\{j : \lambda_j \geq i\}$

Clearly, $\lambda'_1 \geq \lambda'_2 \geq \cdots$

The partition $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ is the *partition conjugate to* λ

Easy homework: Verify that $(\lambda')' = \lambda$ and that $\lambda'_1 = l(\lambda)$

$\mathfrak{S}_m$: The symmetric group on m symbols

The standard action of $\mathfrak{S}_m$ on $\mathbb{R}^m$: For $\tau \in \mathfrak{S}_m$ and $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$, set $\tau \cdot u := (u_{\tau(1)}, \ldots, u_{\tau(m)})$

For each partition $\lambda = (\lambda_1, \ldots, \lambda_m)$ of weight m, define

$$\mathcal{P}_\lambda(f_1, \ldots, f_m) := \prod_{j=1}^{l(\lambda)} \mathbb{E}\Big(\prod_{k=1}^{\lambda_j} f_{\lambda_1 + \cdots + \lambda_{j-1} + k} \Big)$$

Denote by $D(\lambda)$ the set of all $\tau \in \mathfrak{S}_m$ which give rise to *distinct* permutations $\mathcal{P}_\lambda(\tau \cdot (f_1, \ldots, f_m))$ of $\mathcal{P}_\lambda(f_1, \ldots, f_m)$

Conjecture: Let μ be an MTP$_2$ probability measure on L, and $f_1, \ldots, f_m$ be nonnegative increasing functions on L. Then

(i) There exist non-zero constants $\{c_\lambda \in \mathbb{Z} : |\lambda| = m\}$ such that

$$\mathcal{P}_m(f_1, \ldots, f_m) := \sum_{|\lambda|=m} c_\lambda \sum_{\tau \in D(\lambda)} \mathcal{P}_\lambda(\tau \cdot (f_1, \ldots, f_m))$$

is nonnegative

(ii) For $m \geq 3$, there exists a constant d_m such that

$$\mathcal{P}_m(f_1, \ldots, f_{m-1}, 1) \equiv d_m \mathcal{P}_{m-1}(f_1, \ldots, f_{m-1})$$

(iii) If $f_j \equiv 1$, $j = 1, \ldots, m$ then $\mathcal{P}_m(1, \ldots, 1) = 0$; equivalently,

$$\sum_{|\lambda|=m} \operatorname{card}(D(\lambda))\, c_\lambda = 0$$

For $m = 2, 3, 4, 5$, the conjecture is valid

Choose for $\mathcal{P}_m$ the conjugate cumulant κ'_m

For $\lambda = (\lambda_1, \lambda_2, \ldots)$, all coefficients

$$c_\lambda = (-1)^{l(\lambda)-1} \, (\lambda_1 - 1)!$$

in the expansion of $\kappa'_m(f_1, \ldots, f_m)$ are non-zero and satisfy (iii)

For $m = 3, 4, 5$ each κ'_m satisfies (ii) with $d_m = m - 2$

I found evidence that $\kappa'_m(1, \ldots, 1) > 0$, $m \geq 6$

Sahi has constructed a class of functionals τ_m such that

$\tau_m \equiv \kappa'_m$, $m = 2, 3, 4, 5$, and $\tau_m \not\equiv \kappa'_m$, $m \geq 6$

For special MTP$_2$ measures μ, Sahi has proved that $\tau_m \geq 0$ for all m

Sahi explains that there is a crucial difference between the class of partitions of weight ≤ 5 and those of weight > 5

Sahi conjectures that, for $\lambda = (\lambda_1, \lambda_2, \ldots)$, we can take

$$c_\lambda = (-1)^{l(\lambda)-1} (\lambda_1 - 1)! \, (\lambda_2 - 1)! \cdots$$

This differs from my coefficients only if $\lambda_2 \geq 3$, in which case $|\lambda| \geq 6$

Sahi also has obtained FKG inequalities for functions taking values in very general partially ordered algebras, e.g., the algebra of square matrices, symmetric matrices, or polynomials

His proofs are, in a sense, simpler than the original FKG

An amazing coincidence?

Connections with k-characters, group determinants

Ken Johnson's paper

Johnson believes that there is a connection between FKG-type inequalities and Chern classes