7

Correlation

G. R. Brightwell and W. T. Trotter
March 31, 2011

7.1 Introduction

As we saw in the previous chapter, the family $E(\mathbf{P})$ of all linear extensions of the finite poset $\mathbf{P} = (X, P)$ can be thought of as a probability space, with each linear extension given the same probability, namely $1/e(\mathbf{P})$. One can be interested in various events in this probability space: the most fundamental being the events of the form $x \prec y$, for elements $x, y \in X$. If x and y are comparable, this event will have probability 0 or 1, but if they are incomparable then $\Pr(x \prec y)$ will lie strictly between 0 and 1; this is just the fact, from Section 9.99, that there is at least one linear extension of $\mathbf{P}$ with x below y, and one with y below x.

Given various events in a probability space (i.e., subsets of the sample space), it is common to ask for circumstances under which a pair of events is *non-negatively correlated*, i.e., the occurrence of one of the events makes the other no less likely. Equivalently, events A and B in a probability space are *non-negatively correlated* if

$$\Pr(A)\Pr(B) \leq \Pr(A \cap B).$$

If the inequality above is strict, then A and B are *positively correlated*. The definitions of negative and non-positive correlation are analogous.

For instance, it must surely appear intuitively obvious that, for any triple of elements $x, y, z \in X$, the events $x \prec y$ and $x \prec z$ are non-negatively correlated. What this says is that, if we have some information about the underlying linear order on a collection of elements, in the form of a partial order P on X, and we are then given the extra information that element x is below element y, this cannot make it less likely that x is below another element z. Perhaps surprisingly, this "obvious" fact is not so easy to prove, and it spent some time as the "xyz Conjecture" before becoming the "xyz Inequality" when it was proved by Shepp in 1982.

A number of other likely-looking statements about correlation between events in this probability space were proved at around the same time, all using much the same basic techniques, with different ingenious twists. We start this chapter with a look at the main tool, the Ahlswede-Daykin Four Functions Theorem, and then we shall see how it can be used. At the end of the chapter, we turn our attention to a geometric result, the Alexandrov-Fenchel Inequalities, that can be used to prove a slightly different type of result, namely that certain sequences are log-concave.

7.2 The Ahlswede-Daykin Four Functions Theorem

In a 1966 paper [99], Kleitman proved that, if D is a down-set and U an up-set in $\mathbf{2}^n$, then

$$|D|\,|U| \geq 2^n|D \cap U|. \tag{7.1}$$

This is a *correlation inequality*; if one regards $\mathbf{2}^n$ as a probability space, with all elements equally likely, then it says that

$$\Pr(x \in D)\Pr(x \in U) \geq \Pr(x \in D \,\&\, x \in U),$$

in other words, the events "$x \in D$" and "$x \in U$" are non-positively correlated.

Putting $V = \mathbf{2}^n \setminus D$ in Kleitman's theorem yields immediately that

$$|U|\,|V| \leq 2^n|U \cap V| \tag{7.2}$$

for any two up-sets U and V in $\mathbf{2}^n$. This says that the events "$x \in U$" and "$x \in V$" are non-negatively correlated. Indeed, both inequalities are equivalent to: for any up-sets U and V in $\mathbf{2}^n$,

$$|U \cap \overline{V}|\,|V \cap \overline{U}| \leq |U \cap V||\overline{U} \cap \overline{V}|. \tag{7.3}$$

See Exercise 1.

In 1971, Fortuin, Kasteleyn and Ginibre [99], while studying problems in statistical physics, proved the following powerful extension of Kleitman's theorem.

A function μ from a lattice $\mathbf{L}$ to $\mathbb{R}^+$ is said to be *log-supermodular* if $\mu(a)\mu(b) \leq \mu(a \vee b)\mu(a \wedge b)$ for every $a, b \in \mathbf{L}$.

Theorem 7.2.1 (FKG Inequality) *Let* $\mathbf{L}$ *be a distributive lattice, and* *suppose that f, g and μ are functions from $\mathbf{L}$ to $\mathbb{R}^+$ such that f and g are non-decreasing and μ is log-supermodular. Then*

$$\left(\sum_{a\in\mathbf{L}} f(a)\mu(a)\right)\left(\sum_{a\in\mathbf{L}} g(a)\mu(a)\right) \le \left(\sum_{a\in\mathbf{L}} f(a)g(a)\mu(a)\right)\left(\sum_{a\in\mathbf{L}} \mu(a)\right).$$

The FKG Inequality is again a thinly disguised correlation inequality: if μ is normalized so that it becomes a probability measure, Theorem 7.2.1 says simply that $(\mathbb{E}f)(\mathbb{E}g) \le \mathbb{E}(fg)$. In many, but not all, applications, one takes μ to be constant. If f and g are the indicator functions of up-sets U and V, and $\mu = 1$, then the FKG Inequality gives the following result.

Corollary 7.2.2 *For any two up-sets U, V in a distributive lattice $\mathbf{L}$,*

$$|U|\,|V| \le |U \cap V|\,|\mathbf{L}|.$$

This result says that any two up-sets in a distributive lattice are non-negatively correlated; as before, one can deduce that any two down-sets are non-negatively correlated, and that an up-set and a down-set are non-positively correlated. For $L = \mathbf{2}^n$, Corollary 7.2.2 is of course just (7.2), equivalent to Kleitman's result.

The FKG Inequality is a beautiful result, and for many an application it is all one really needs. Indeed, most of our correlation results use only Corollary 7.2.2. However, in 1978 Ahlswede and Daykin gave a result which includes this and many other inequalities, and whose proof is at least as simple as any of the earlier ones. The statement is a little more involved, and it is perhaps something of a surprise that the full power of the theorem is sometimes needed to derive important results.

For S any finite set, $f : S \to \mathbb{R}^+$ a function, and $A \subseteq S$, we adopt the usual and useful convention that $f(A)$ denotes $\sum_{a\in A} f(a)$. Also, for A and B subsets of a lattice $\mathbf{L}$, set $A \vee B = \{a \vee b : a \in A, b \in B\}$ and $A \wedge B = \{a \wedge b : a \in A, b \in B\}$. If A and B are up-sets, then $A \vee B$ is just $A \cap B$ (Exercise 7).

Theorem 7.2.3 (Ahlswede-Daykin Four Functions Theorem) *Let $\mathbf{L}$ be a finite distributive lattice, and let α, β, γ and δ be four functions from $\mathbf{L}$ to $\mathbb{R}^+$ such that $\alpha(a)\beta(b) \le \gamma(a \vee b)\delta(a \wedge b)$ for all $a, b \in \mathbf{L}$. Then, for any subsets $A, B \subseteq \mathbf{L}$,*

$$\alpha(A)\beta(B) \le \gamma(A \vee B)\delta(A \wedge B).$$

Let us see how we can use the Four Functions Theorem to prove the FKG Inequality. Given non-decreasing functions f and g, and a log-supermodular function μ, all from a distributive lattice $\mathbf{L}$ to the non-negative reals, set $A = B = \mathbf{L}$, $\alpha = f\mu$, $\beta = g\mu$, $\gamma = fg\mu$ and $\delta = \mu$. Clearly $A \vee B = A \wedge B = \mathbf{L}$, and so the conclusion of the Four Functions Theorem is exactly what we want, so it remains to check that

$$f(a)\mu(a)g(b)\mu(b) \leq f(a \vee b)g(a \vee b)\mu(a \vee b)\mu(a \wedge b),$$

but this is immediate from the assumptions on f, g and μ.

If α, β, γ and δ are all identically 1, the Four Functions Theorem gives us the following corollary, extending Kleitman's Theorem.

Corollary 7.2.4 *For A and B subsets of a distributive lattice $\mathbf{L}$,*

$$|A|\,|B| \leq |A \vee B|\,|A \wedge B|.$$

Corollary 7.2.4 had earlier been proved by Daykin [99], who also observed that the inequality fails in any non-distributive lattice: see Exercise 6. If A and B are up-sets, then we obtain a slightly strengthened version of Corollary 7.2.2.

We shan't prove the Four Functions Theorem, but let us at least indicate how the proof goes. Firstly, we observe that it is enough to prove the theorem for $A = B = \mathbf{L}$. Indeed, if we have the result in this case, and we are given sets $A, B \subseteq \mathbf{L}$, and four functions $\alpha, \beta, \gamma, \delta$ satisfying $\alpha(a)\beta(b) \leq \gamma(a \vee b)\delta(a \wedge b)$, then set:

$$\begin{aligned} \alpha'(a) &= \alpha(a)\chi(a \in A), \\ \beta'(a) &= \beta(a)\chi(a \in B), \\ \gamma'(a) &= \gamma(a)\chi(a \in A \vee B), \\ \delta'(a) &= \delta(a)\chi(a \in A \wedge B). \end{aligned}$$

These functions satisfy $\alpha'(a)\beta'(b) \leq \gamma'(a \vee b)\delta'(a \wedge b)$, and applying our special case of the Four Functions Theorem gives us $\alpha(A)\beta(B) \leq \gamma(A \vee B)\delta(A \wedge B)$, as we require. Furthermore, since every distributive lattice is isomorphic to a sublattice of $\mathbf{2}^n$ for some n (by Theorem 9.99), the same technique shows that it is enough to prove the result for $\mathbf{L} = \mathbf{2}^n$. Now one uses induction on n: the case $n = 1$ requires a little detailed analysis, and this same analysis also forms the heart of the induction step. See the original paper of Ahlswede and Daykin [99] or, for instance, Bollobás [99] for details.

7.3 Order-preserving Maps

We now begin to apply the results of the previous section to problems of interest to us. Our line of attack will be to apply the Four Functions Theorem (perhaps not in full generality). For this we need to get our hands on a distributive lattice $\mathbf{L}$ connected to the problem, such that summing various functions over $\mathbf{L}$ will produce the quantities we are interested in.

For instance, we want to prove the xyz inequality described in Section 7.1. It would be ideal if we could find a distributive lattice whose elements are exactly the linear extensions of our poset $\mathbf{P}$, where the set of linear extensions with x below y forms an up-set, as does the set with x below z. Sadly, this does not appear to be possible for an arbitrary poset $\mathbf{P}$. However, we can do almost as well if, instead of working with linear extensions directly, we work with order-preserving maps from $\mathbf{P}$ to $[k]$, where k is a (large) positive integer. The idea is that the set of order-preserving maps *can* be made into a distributive lattice in various useful ways, and that any correlation results we obtain in the probability space of order-preserving maps can be carried over into the space of linear extensions. This approach, and the results and proofs derived using it, are all due to Shepp [99].

For $\mathbf{P} = (X, P)$ a finite poset and $k \in \mathbb{N}$, let $M_k(\mathbf{P})$ denote the set of all order-preserving maps from $\mathbf{P}$ to $[k]$, i.e., the set of functions $\omega : X \to \{1, \ldots, k\}$ such that $\omega(x) \preceq \omega(y)$ whenever $(x, y) \in P$. We make $M_k(\mathbf{P})$ into a probability space by declaring all of its elements to be equally likely, and denote the probability measure in this space by $\Pr_k$. Note that, if $\mathbf{Q}$ is an extension of $\mathbf{P}$, then $M_k(\mathbf{Q}) \subseteq M_k(\mathbf{P})$. Also, $M_k(\mathbf{P}) = \bigcup_{\mathbf{R} \in E(\mathbf{P})} M_k(\mathbf{R})$; however, this is not a disjoint union – indeed, all the constant functions ω are in all the $M_k(\mathbf{R})$. If A is any subset of $E(\mathbf{P})$ (i.e., any event in the probability space $E(\mathbf{P})$), set $A_k = \bigcup_{\mathbf{R} \in A} M_k(\mathbf{R})$, an event in the probability space $M_k(\mathbf{P})$. The basic lemma that drives the approach is as follows.

Lemma 7.3.1 *For any finite poset $\mathbf{P} = (X, P)$, and any subset A of $E(\mathbf{P})$,*

$$\Pr_k(A_k) \to \frac{|A|}{e(\mathbf{P})}$$

as $k \to \infty$.

Proof Set $n = |X|$, as usual. Let $b_k(\mathbf{P})$ be the probability, in $M_k(\mathbf{P})$, that there is some pair of elements x, y with $\omega(x) = \omega(y)$. It is intuitively clear, and easy to check (see Exercise 11), that $b_k(\mathbf{P}) \to 0$ as $k \to \infty$.

Now, among those order-preserving maps $\omega \in M_k(\mathbf{P})$ for which all the images $\omega(x)$ are distinct, there are exactly $\binom{k}{n}$ in $M_k(\mathbf{R})$, for each linear extension $\mathbf{R}$ of $\mathbf{P}$. Thus, for any event $A \subseteq E(\mathbf{P})$, the probability of A_k, conditioned on the event that all the $\omega(x)$ are distinct, is exactly $|A|/e(\mathbf{P})$.

Thus we have

$$(1 - b_k(\mathbf{P}))\frac{|A|}{e(\mathbf{P})} \leq \Pr_k(A_k) \leq (1 - b_k(\mathbf{P}))\frac{|A|}{e(\mathbf{P})} + b_k(\mathbf{P}),$$

and the result follows. □

We immediately obtain the following corollary.

Corollary 7.3.2 *Let $\mathbf{P} = (X, P)$ be a finite poset, and let A and B be subsets of $E(\mathbf{P})$. Suppose that, for sufficiently large k, the events A_k and B_k are non-negatively correlated in $M_k(\mathbf{P})$. Then A and B are non-negatively correlated in $E(\mathbf{P})$.*

We now want to impose a distributive lattice structure on $M_k(\mathbf{P})$ such that events of interest to us become up-sets or down-sets. One obvious possibility is to define, for $\omega, \nu \in M_k(\mathbf{P})$, the meet $\omega \wedge \nu$ and join $\omega \vee \nu$ by

$$\begin{aligned} (\omega \wedge \nu)(x) &= \min(\omega(x), \nu(x)), \\ (\omega \vee \nu)(x) &= \max(\omega(x), \nu(x)). \end{aligned}$$

It is easy to check that this does define a distributive lattice, but unfortunately no event of the form $\omega(x) \prec \omega(y)$ is an up-set or down-set in the lattice. One *can* apply any of the correlation inequalities from Section 7.2 to show that, for instance, for any two elements $x, y \in X$, $\omega(x)$ and $\omega(y)$ are non-negatively correlated random variables. But this is not what we are looking for.

However, our first correlation result for partial orders, proved by Shepp [99], uses a lattice structure on $M_k(\mathbf{P})$ that is only slightly more complicated.

For a finite poset $\mathbf{P} = (X, P)$ and any subset S of $X \times X$, let G_S denote the event, in $E(\mathbf{P})$, that $x \prec y$ for every $(x, y) \in S$. Similarly, let $G_{k,S}$ denote the event, in $M_k(\mathbf{P})$, that $\omega(x) \preceq \omega(y)$ whenever $(x, y) \in S$.

Theorem 7.3.3 *Let $\mathbf{P} = (X, P)$ be a finite poset such that X is the disjoint union of Y and Z, with $Y|Z$, and let S and T be two subsets of $Y \times Z$. For any $k \in \mathbb{N}$, the events $G_{k,S}$ and $G_{k,T}$ are non-negatively correlated in $M_k(\mathbf{P})$, and the events G_S and G_T are non-negatively correlated in $E(\mathbf{P})$.*

In particular, Theorem 7.3.3 says that, if Y and Z are as above, with $y_1, y_2 \in Y$ and $z_1, z_2 \in Z$, then the events $y_1 \prec z_1$ and $y_2 \prec z_2$ are non-negatively correlated. One way to look at this is to think of Y and Z as two "teams". Some information may be known in advance about the relative ranking of players within each team, but nothing else. Now, the information that player y_1 on team Y is ranked below player z_1 on team Z makes it no less likely that y_2 is ranked below z_2.

Proof By Corollary 7.3.2, it is enough to prove the result for $M_k(\mathbf{P})$. We give $M_k(\mathbf{P})$ the following lattice structure. For $\omega, \nu \in M_k(\mathbf{P})$, set:

$$\begin{aligned}
(\omega \wedge \nu)(y) &= \min(\omega(y), \nu(y)) && \text{for } y \in Y,\\
(\omega \wedge \nu)(z) &= \max(\omega(z), \nu(z)) && \text{for } z \in Z,\\
(\omega \vee \nu)(y) &= \max(\omega(y), \nu(y)) && \text{for } y \in Y,\\
(\omega \vee \nu)(z) &= \min(\omega(z), \nu(z)) && \text{for } z \in Z.
\end{aligned}$$

So, in the underlying partial order, $\omega \le \nu$ if and only if $\omega(y) \preceq \nu(y)$ for all $y \in Y$ and $\omega(z) \succeq \nu(z)$ for all $z \in Z$.

To see that this defines a distributive lattice, we note first that it is a subposet of the poset of *all* functions from X to $[k]$, with the partial order given above, and this is a cartesian product of chains and so is a distributive lattice. Thus all we need to check is that $M_k(\mathbf{P})$ is closed under the operations $\wedge$ and $\vee$ defined above. Certainly, for $\omega, \nu \in M_k(\mathbf{P})$, $\omega \wedge \nu$ and $\omega \vee \nu$ are maps from X to $[k]$. Now, if $y_1 < y_2$, then $\omega(y_1) \preceq \omega(y_2)$ and $\nu(y_1) \preceq \nu(y_2)$, so $\min(\omega(y_1), \nu(y_1)) \preceq \min(\omega(y_2), \nu(y_2))$ and $\max(\omega(y_1), \nu(y_1)) \preceq \max(\omega(y_2), \nu(y_2))$, so $\omega \wedge \nu$ and $\omega \vee \nu$ preserve the order on Y, and, by symmetry, also the order on Z. Since $Y|Z$ in $\mathbf{P}$, we are done.

Now observe that, in this distributive lattice, the events $G_{k,S}$ and $G_{k,T}$ are down-sets. Thus, by (for instance) Corollary 7.2.2, these two events are non-negatively correlated, as required. □

Theorem 7.3.3 can obviously be extended to cover pairs of more complex events, each a disjunction of events of the form G_S, with $S \subseteq Y \times Z$.

Theorem 7.3.3 is not particularly impressive at first sight: one is tempted to think that there should be a short direct proof, without needing to resort to "trickery". However, we know of no such proof. In fact, the result is surprisingly useful. Here, for instance, is a consequence, due to Fishburn [99], which we shall use later.

Theorem 7.3.4 *Let U and V be up-sets in a finite poset $\mathbf{P} = (X, P)$. Then*

$$\frac{e(\mathbf{P}_U)e(\mathbf{P}_V)}{e(\mathbf{P}_{U\cap V})e(\mathbf{P}_{U\cup V})} \le \frac{|U|!\,|V|!}{|U\cap V|!\,|U\cup V|!} \le 1.$$

Note that, for instance, $e(\mathbf{P}_U)/|U|!$ is the probability, in the space of all linear orders on $U \cup V$, of G_{P_U}. As $G_{P_{U\cup V}} = G_{P_U} \cap G_{P_V}$—every pair of

Fig. 7.1. The various sets in Theorem 7.3.4

elements comparable in $P_{U \cup V}$ is contained in either U or V, as U and V are up-sets—Theorem 7.3.4 says exactly that the events G_{P_U} and G_{P_V} are non-negatively correlated in the space of all linear orders on X.

Proof Let $Z = U \cap V$ and $Y = U \triangle V$. Note that, in $\mathbf{P}$, $(U \setminus V) | (V \setminus U)$.

The basic idea is that $P_{U \cup V}$ can be decomposed into independent partial orders on Z, $U \setminus V$ and $V \setminus U$, together with the two sets $S = P \cap [(U \setminus V) \times Z]$ and $T = P \cap [(V \setminus U) \times Z]$. The theorem of Shepp then tells us that G_S and G_T are non-negatively correlated, which turns out to be exactly what we need. See Figure 7.1.

More formally, we define a partial order $\mathbf{Q} = (W, Q)$ on $W = U \cup V = Y \cup Z$ to be the disjoint union of $\mathbf{P}_Z$ and $\mathbf{P}_Y$, so $Q = P \cap [(Z \times Z) \cup (Y \times Y)]$. Now we apply Shepp's Theorem 7.3.3 to the poset $\mathbf{Q}$, with partition $W = Y \cup Z$, and the events G_S and G_T, obtaining that $\Pr(G_S) \Pr(G_T) \leq \Pr(G_S \,\&\, G_T)$, in $E(\mathbf{Q})$.

In calculating $\Pr(G_S)$, note that the elements of $V \setminus U$ can be ignored, and that $Q_U \cup S = P_U$, so that

$$\Pr(G_S) = \frac{e(\mathbf{P}_U)}{e(\mathbf{Q}_U)} = \frac{e(\mathbf{P}_U)}{e(\mathbf{P}_{U \cap V}) e(\mathbf{P}_{U \setminus V}) \binom{|U|}{|U \cap V|}}.$$

Similarly

$$\Pr(G_T) = \frac{e(\mathbf{P}_V)}{e(\mathbf{Q}_V)} = \frac{e(\mathbf{P}_V)}{e(\mathbf{P}_{U \cap V}) e(\mathbf{P}_{V \setminus U}) \binom{|V|}{|U \cap V|}},$$

and

$$\begin{aligned}\Pr(G_S \,\&\, G_T) &= \frac{e(\mathbf{P}_{U\cup V})}{e(\mathbf{Q}_{U\cup V})}\\ &= \frac{e(\mathbf{P}_{U\cup V})}{e(\mathbf{P}_{U\cap V})e(\mathbf{P}_{U\setminus V})e(\mathbf{P}_{V\setminus U})\binom{|U\cup V|}{|U\cap V|,|U\setminus V|,|V\setminus U|}}.\end{aligned}$$

Rearranging terms now gives the required inequality. □

To prove the xyz inequality, Shepp [99] required another, again slightly more complicated, lattice structure on $M_k(\mathbf{P})$.

Theorem 7.3.5 (xyz Inequality) *Let x, y and z be any three distinct elements of a finite poset $\mathbf{P} = (X, P)$. Then the events $\omega(x) \preceq \omega(y)$ and $\omega(x) \preceq \omega(z)$ are non-negatively correlated in $M_k(\mathbf{P})$. Also, the events $x \prec y$ and $x \prec z$ are non-negatively correlated in $E(\mathbf{P})$.*

Proof Again, Corollary 7.3.2 tells us that we need only prove the result for $M_k(\mathbf{P})$.

For convenience, we label the elements of X as $x = x_1, y = x_2, z = x_3, x_4, \ldots, x_n$. Now consider the function Φ from $M_k(\mathbf{P})$ to $\mathcal{H}_k = [-k, -1] \times [-k+1, k-1] \times [-k+1, k-1] \times \cdots \times [-k+1, k-1]$ given by:

$$\Phi(\omega) = \big(-\omega(x_1), \omega(x_2) - \omega(x_1), \omega(x_3) - \omega(x_1), \ldots, \omega(x_n) - \omega(x_1)\big).$$

The map Φ is clearly an injection; we claim that $\Phi(M_k(\mathbf{P}))$ is a sublattice of the distributive lattice $\mathcal{H}_k$. This allows us to pass the distributive lattice structure on to $M_k(\mathbf{P})$. Clearly then the events $\omega(x_j) - \omega(x_1) \succeq 0$ are all up-sets in this lattice, so non-negatively correlated by Corollary 7.2.2. For $j = 2, 3$, the events in question are exactly the events $\omega(x) \preceq \omega(y)$ and $\omega(x) \preceq \omega(z)$.

It remains to verify that $\Phi(M_k(\mathbf{P}))$ is a sublattice of $\mathcal{H}_k$. Suppose then that ω and ν are two elements of $M_k(\mathbf{P})$, and consider $\Phi(\omega) \wedge \Phi(\nu)$. The first co-ordinate of this vector is $-\max(\omega(x_1), \nu(x_1))$, and for $j > 1$ the jth co-ordinate is $\min(\omega(x_j) - \omega(x_1), \nu(x_j) - \nu(x_1))$. Accordingly, we define μ by $\mu(x_1) = \max(\omega(x_1), \nu(x_1))$ and, for $j > 1$,

$$\mu(x_j) = \mu(x_1) + \min(\omega(x_j) - \omega(x_1), \nu(x_j) - \nu(x_1)).$$

It is easy to check that

$$\min(\omega(x_j), \nu(x_j)) \le \mu(x_j) \le \max(\omega(x_j), \mu(x_j)),$$

so $\mu(x_j) \in [k]$; we need also to check that μ is order-preserving. If $x_j < x_\ell$ in $\mathbf{P}$, with $j, \ell > 1$, then $\omega(x_j) - \omega(x_1) \preceq \omega(x_\ell) - \omega(x_1)$ and $\nu(x_j) - \nu(x_1) \preceq$

$\nu(x_\ell) - \nu(x_1)$, so the same also holds for μ, and $\mu(x_j) \preceq \mu(x_\ell)$. Also, if x_j is less than (greater than) x_1, then both $\omega(x_j) - \omega(x_1)$ and $\nu(x_j) - \nu(x_1)$ are non-positive (non-negative), and hence so is $\mu(x_j) - \mu(x_1)$. Thus μ is indeed in $M_k(\mathbf{P})$, and so $\Phi(M_k(\mathbf{P}))$ is closed under meets. In exactly the same way, one verifies that $\Phi(M_k(\mathbf{P}))$ is closed under joins, so it is a sublattice of $\mathcal{H}_k$, as claimed.

This completes the proof. □

Again, one can prove slightly more general correlation inequalities by working in the same lattice; if x is a fixed element of a poset, and A and B are any events made up from basic events of the form $x \prec x_j$ by conjunctions and disjunctions, then A and B are non-negatively correlated.

In the next section, we shall give another proof of the xyz Inequality, showing also that, except in trivial cases, the inequality holds strictly.

7.4 Linear Extensions

In the previous section, we looked at ways of proving correlation results in $E(\mathbf{P})$ via the space $M_k(\mathbf{P})$ of order-preserving maps; in this section, we proceed more directly.

We begin with a result which can be viewed as a companion to Theorem 7.3.3. Again, suppose we have a scenario with two "teams" Y and Z. We are interested in conditions implying that events $y_i \prec z_i$ ($y_i \in Y$, $z_i \in Z$) are non-negatively correlated in $E(\mathbf{P})$. Roughly speaking, we need our condition to tie pairs of elements from the same team more closely than pairs from opposite teams. In Theorem 7.3.3, this was accomplished by assuming that we had no information about rankings between the teams. The other extreme, where we have *complete* information about the ranking inside each team, is what we consider now. The following result was first proved by Graham, Yao and Yao [99]; the proof here is due to Shepp [99].

Theorem 7.4.1 *Suppose that* $\mathbf{P} = (X, P)$ *is a poset of width 2, with a chain-partition* (Y, Z)*, and let* S *and* T *be subsets of* $Y \times Z$*. Then* G_S *and* G_T *are non-negatively correlated in* $E(\mathbf{P})$*.*

Proof We label the elements of Y so that $y_1 < y_2 < \cdots < y_\ell$ and the elements of Z so that $z_1 < z_2 < \cdots < z_m$. We would like to put a distributive lattice structure on $E(\mathbf{P})$ so that events of the form $y_i \prec z_j$ are down-sets. Since all we need to do to determine a linear extension of $\mathbf{P}$ is specify which events of this form occur in the particular linear extension, a natural possibility is to look at the subset lattice $\mathcal{P}(K)$, where $K = \{(y, z) : y \in Y, z \in Z, y|z\}$.

The set $E(\mathbf{P})$ can be thought of as a subset of $\mathcal{P}(K)$; we claim that it is a sublattice.

Indeed, which subsets J of K correspond to linear extensions of $\mathbf{P}$? A necessary condition is that, for each $y_i \in Y$, the set $\{z_j : (y_i, z_j) \in J\}$ is an up-set in the chain Z and, for each $z_j \in Z$, the set $\{y_i : (y_i, z_j) \in J\}$ is a down-set in the chain Y. This condition is also sufficient since it implies transitivity in $P \cup J \cup \{(z, y) : (y, z) \notin J\}$.

Now, given any two subsets J_1 and J_2 of K satisfying the above condition, the union $J_1 \cup J_2$ and intersection $J_1 \cap J_2$ also satisfy the condition. Thus $E(\mathbf{P})$ corresponds to a sublattice of the distributive lattice $\mathcal{P}(K)$.

It remains only to observe that G_S and G_T are up-sets in this lattice, so non-negatively correlated by Corollary 7.2.2. □

Now we come to Fishburn's proof [99] of the *strict xyz* inequality, stating that, except in trivial cases, the xyz inequality holds strictly. These "trivial cases" are as follows.

- If either (x, y) or (x, z) is in P, then one of $\Pr(x \prec y)$ and $\Pr(x \prec z)$ is equal to 1 and the other is equal to $\Pr(x \prec y \,\&\, x \prec z)$.
- If either (y, x) or (z, x) is in P, then both $\Pr(x \prec y) \Pr(x \prec z)$ and $\Pr(x \prec y \,\&\, x \prec z)$ are equal to 0.

On the other hand, if x is incomparable to both y and z in P, but $(y, z) \in P$, then $\Pr(x \prec y) = \Pr(x \prec y \,\&\, x \prec z) > 0$, but $\Pr(x \prec z) < 1$, so we have strict inequality. Thus we may assume that x, y and z form a 3-element antichain.

Observe that the inequality $\Pr(x \prec y) \Pr(x \prec z) < \Pr(x \prec y \,\&\, x \prec z)$ can be written as

$$\Pr(z \prec x \prec y) \Pr(y \prec x \prec z) < \Pr(x \prec y, z) \Pr(y, z \prec x); \tag{7.4}$$

this is an instance of the easy result in Exercise 1.

Fishburn [99] not only proved this inequality, but also found the maximum value of the ratio of the two sides, as a function of $n = |X|$.

Theorem 7.4.2 *Let x, y and z be three mutually incomparable elements in an n-element poset $\mathbf{P} = (X, P)$. Then*

$$\Pr(z \prec x \prec y) \Pr(y \prec x \prec z) \le \eta_n \Pr(x \prec y, z) \Pr(y, z \prec x),$$

where $\eta_n = \left(\frac{n-1}{n+1}\right)^2$ if n is odd and $\eta_n = \frac{n-2}{n+2}$ if n is even.

Proof As usual, the first step is to construct a suitable distributive lattice. We break $X \setminus \{x\}$ into three pieces: $U = U(x)$, $D = D(x)$ and $I = I(x)$, being the sets of elements above, below and incomparable with x in $\mathbf{P}$, respectively. So y and z are in I. Now, in each linear extension λ of $\mathbf{P}$, consider the set of elements above x; this consists of $U(x)$, together with some up-set $A(\lambda)$ in $\mathbf{P}_I$. This suggests the approach of using the set of up-sets in $\mathbf{P}_I$, ordered by containment, as our distributive lattice.

For a subset A of I, set $\overline{A} = I \setminus A$. For any up-set A of I, let $\mu(A)$ be the number of linear extensions λ of $\mathbf{P}$ with $A(\lambda) = A$. We have that $\mu(A) = e(\mathbf{P}_{U\cup A})e(\mathbf{P}_{D\cup\overline{A}})$. If A and B are two up-sets, then we have

$$\frac{\mu(A)\mu(B)}{\mu(A\cup B)\mu(A\cap B)} = \frac{e(\mathbf{P}_{U\cup A})e(\mathbf{P}_{U\cup B})}{e(\mathbf{P}_{U\cup A\cup B})e(\mathbf{P}_{U\cup(A\cap B)})}\frac{e(\mathbf{P}_{D\cup\overline{A}})e(\mathbf{P}_{D\cup\overline{B}})}{e(\mathbf{P}_{D\cup(\overline{A}\cap\overline{B})})e(\mathbf{P}_{D\cup\overline{A}\cup\overline{B}})}.$$

Theorem 7.3.4, or its dual, applies to both fractions on the right above, and we obtain that

$$\frac{\mu(A)\mu(B)}{\mu(A\cup B)\mu(A\cap B)} \le \frac{|U\cup A|!\,|U\cup B|!}{|U\cup A\cup B|!\,|U\cup(A\cap B)|!}\frac{|D\cup\overline{A}|!\,|D\cup\overline{B}|!}{|D\cup(\overline{A}\cap\overline{B})|!\,|D\cup\overline{A}\cup\overline{B}|!}. \tag{7.5}$$

If our sole aim was to prove the xyz inequality, we would now be done, since we could use the FKG Inequality with this choice of μ—the right hand side of (7.5) is at most 1, so μ is log-supermodular—with $f(A) = \chi(y \in A)$, and with $g(A) = \chi(z \in A)$. For instance, $\sum_A f(A)\mu(A)$ is the total number of linear extensions λ of $\mathbf{P}$ with $y \in A(\lambda)$, i.e., the total number of linear extensions with y above x. To take advantage of the stronger condition satisfied by μ, we will need to use the greater flexibility provided by the Four Functions Theorem.

To begin with, we investigate how close the ratio of factorials in (7.5) can be to 1. If $A \subseteq B$ or $B \subseteq A$, then the ratio is identically 1, but we are interested in the case where neither $A \setminus B$ nor $B \setminus A$ is empty. Suppose without loss of generality that $|U \cup A| = m$, and $|U \cup B| = m + k$, with $k \ge 0$, so that $|D \cup \overline{A}| = n - m - 1$ and $|D \cup \overline{B}| = n - m - k - 1$. Subject to this, both $|U \cup A \cup B|!\,|U \cup (A \cap B)|!$ and $|D \cup (\overline{A} \cap \overline{B})|!\,|D \cup \overline{A} \cup \overline{B}|!$ are minimized when $A \setminus B$ consists of a single element, and $|B \setminus A| = k + 1$. In this case,

$$\frac{|U\cup A|!\,|U\cup B|!}{|U\cup A\cup B|!\,|U\cup(A\cap B)|!}\;\frac{|D\cup\overline{A}|!\,|D\cup\overline{B}|!}{|D\cup(\overline{A}\cap\overline{B})|!\,|D\cup\overline{A}\cup\overline{B}|!} = \frac{m}{m+k+1}\frac{n-m-k-1}{n-m},$$

which is maximized when $k = 0$ and m is as close to $(n-1)/2$ as possible, and in this case the ratio is equal to η_n.

Thus we have that, if neither $A \setminus B$ nor $B \setminus A$ is empty, then

$$\mu(A)\mu(B) \leq \eta_n \mu(A \cup B)\mu(A \cap B). \tag{7.6}$$

Now we apply the Four Functions Theorem, Theorem 7.2.3. The lattice $\mathbf{L}$ we use is the up-set lattice of $\mathbf{P}_I$, and we set

$$\begin{aligned}
\alpha(A) &= \mu(A)\chi(y \in A, z \notin A)\\
\beta(A) &= \mu(A)\chi(z \in A, y \notin A)\\
\gamma(A) &= \mu(A)\chi(y, z \in A)\\
\delta(A) &= \eta_n \mu(A)\chi(y, z \notin A).
\end{aligned}$$

To apply the Four Functions Theorem, we need to check that, for any up-sets A and B in $\mathbf{P}_I$,

$$\alpha(A)\beta(B) \leq \gamma(A \cup B)\delta(A \cap B).$$

The left-hand side is zero unless $y \in A \setminus B$ and $z \in B \setminus A$; in that case, the left-hand side is $\mu(A)\mu(B)$, the right-hand side is $\eta_n \mu(A \cup B)\mu(A \cap B)$, and the inequality follows from (7.6).

We deduce that

$$\alpha(\mathbf{L})\beta(\mathbf{L}) \leq \gamma(\mathbf{L})\delta(\mathbf{L}).$$

Now, for instance, $\alpha(\mathbf{L})$ is the total number of linear extensions λ with $y \in A(\lambda)$, $z \notin A(\lambda)$, i.e., the total number with z below x and y above. On dividing each term by $e(\mathbf{P})$, we obtain

$$\Pr(z \prec x \prec y)\Pr(y \prec x \prec z) \leq \eta_n \Pr(x \prec y, z)\Pr(y, z \prec x),$$

as required. □

To see that Theorem 7.4.2 is best possible, consider the poset formed by taking a chain of $n-2$ elements, with x as close to the center as possible, together with isolated elements y and z. If n is odd, then there are $(n-3)/2$ elements above and below x on the chain, and we have $\Pr(z \prec x \prec y) = \Pr(y \prec x \prec z) = \frac{1}{4}(n-1)^2$, while $\Pr(x \prec y, z) = \Pr(y, z \prec x) = \frac{1}{4}(n-1)(n+1)$, so we have

$$\Pr(z \prec x \prec y)\Pr(y \prec x \prec z) = \left(\frac{n-1}{n+1}\right)^2 \Pr(x \prec y, z)\Pr(y, z \prec x),$$

and the fraction here is just η_n. The calculation for the case of n even is similar.

7.5 Sequences in posets, and the Alexandrov-Fenchel inequalities

Besides the Four Functions Theorem and its relatives, there is one other tool that has proved useful for proving correlation inequalities of a different style, going by the name of the *Alexandrov-Fenchel Inequalities for Mixed Volumes.*

This tool can be used to prove that certain sequences $z_0, z_1, \ldots, z_m$ are *log-concave*: $z_{i-1}z_{i+1} \le z_i^2$, for $i = 1, \ldots, m-1$. A log-concave sequence z_i is necessarily *unimodular*: increasing up to a certain point and decreasing thereafter.

In this section, we shall present the Alexandrov-Fenchel Inequalities in conjunction with one of the key applications, given by Stanley [99] in an influential paper of 1981.

For $\mathbf{P} = (X, P)$ an n-element poset, $x \in X$, and $\prec$ a linear extension of P, let $h_\prec(x)$ denote the *height* of x in $\prec$, i.e., the number of elements below or equal to x in $\prec$, so $h_\prec(x) \in [n]$. Now, for $h = 1, \ldots, n$, let $e(\mathbf{P}; x \to h)$ denote the number of linear extensions $\prec$ of P in which $h_\prec(x) = h$. Stanley proved that the sequence $e(\mathbf{P}; x \to h)$, $h = 1, \ldots, n$, is log-concave (and therefore unimodular).

The fact that this sequence is unimodular hopefully accords with intuition: an element x has a "most probable height" h, and the probability that its height takes a value h' drops off as h' moves away from h.

Suppose for example that x is only incomparable with the two elements y and z; then $h_\prec(x)$ takes just three values $i-1$, i and $i+1$, say. The log-concavity of $e(\mathbf{P}; x \to h)$ is equivalent to:

$$(\Pr(z \prec x \prec y) + \Pr(y \prec x \prec z))^2 \ge \Pr(x \prec y, z) \Pr(y, z \prec x).$$

Compare this with (7.4), a statement equivalent to the xyz inequality, and we see that this is in some sense a negative correlation equality.

We will approach Stanley's proof indirectly. We saw in the previous chapter that the volume of the order polytope $\mathcal{O}(\mathbf{P})$ is equal to $e(\mathbf{P})/n!$, for $\mathbf{P} = (X, P)$ an n-element poset. We fix an element x of X and a constant $\lambda \in [0,1]$, and ask for the $(n-1)$-dimensional volume of the *slice*

$$\mathcal{O}(\mathbf{P}; x \to \lambda) = \{\mathbf{a} \in \mathcal{O}(\mathbf{P}) : a_x = \lambda\}.$$

We calculate this in two ways.

First, let's follow the simple argument we used before, and break $\mathcal{O}(\mathbf{P})$ into $e(\mathbf{P})$ pieces, according to the order of the coordinates. For a linear extension $\prec$ of P, in which $h_\prec(x) = h$, the $(n-1)$-dimensional volume of the slice $\mathcal{O}(\prec; x \to \lambda)$ is given by

$$\frac{\lambda^{h-1}}{(h-1)!}\frac{(1-\lambda)^{n-h}}{(n-h)!}.$$

To see this, note first that the volume of that part of $[0,1]^X$ in which $a_y \leq a_x = \lambda$ for $y \prec x$ and $a_y \geq a_x = \lambda$ for $x \prec y$ is $\lambda^{h-1}(1-\lambda)^{n-h}$. This region is partitioned into $(h-1)!\,(n-h)!$ pieces of equal volume, according to the order of the coordinates, and one of these corresponds to the slice $\mathcal{O}(\prec; x \to \lambda)$.

Therefore, summing over all linear extensions, we have

$$\mathrm{Vol}_{n-1}(\mathcal{O}(\mathbf{P}; x \to \lambda)) = \sum_{h=1}^{n} e(\mathbf{P}; x \to h)\frac{\lambda^{h-1}}{(h-1)!}\frac{(1-\lambda)^{n-h}}{(n-h)!}, \qquad (7.7)$$

where Vol_{n-1} denotes $(n-1)$-dimensional volume.

Let $\mathbf{P}(x;0)$ denote the poset obtained from $\mathbf{P}$ by deleting $D[x]$, and $\mathbf{P}(x;1)$ the poset obtained by deleting $U[x]$. The order polytope $\mathcal{B} = \mathcal{O}(\mathbf{P}(x;0))$ is exactly the slice $\mathcal{O}(\mathbf{P}; x \to 0)$ and similarly $\mathcal{A} = \mathcal{O}(\mathbf{P}(x;1)) = \mathcal{O}(\mathbf{P}; x \to 1)$. For $\lambda \in [0,1]$, consider the *Minkowski sum* $\lambda\mathcal{A} + (1-\lambda)\mathcal{B} = \{\lambda\mathbf{a} + (1-\lambda)\mathbf{b} : \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$. We claim that this is exactly $\mathcal{O}(\mathbf{P}; x \to \lambda)$.

It's clear that $\lambda\mathcal{A} + (1-\lambda)\mathcal{B} \subseteq \mathcal{O}(\mathbf{P}; x \to \lambda)$. For the converse, let $\mathbf{c}$ be a vector in $\mathcal{O}(\mathbf{P}; x \to \lambda)$ and define, for $y \in X$,

$$a_y = \begin{cases} c_y/\lambda & c_y \leq \lambda \\ 1 & c_y \geq \lambda; \end{cases} \qquad b_y = \begin{cases} 0 & c_y \leq \lambda \\ (c_y - \lambda)/(1-\lambda) & c_y \geq \lambda. \end{cases}$$

Now it is evident that $\mathbf{a} \in \mathcal{A}$, $\mathbf{b} \in \mathcal{B}$, and $\lambda\mathbf{a} + (1-\lambda)\mathbf{b} = \mathbf{c}$. Thus $\mathcal{O}(\mathbf{P}; x \to \lambda) = \lambda\mathcal{A} + (1-\lambda)\mathcal{B}$.

Combining this with 7.7 gives us that

$$\mathrm{Vol}_{n-1}(\lambda\mathcal{A} + (1-\lambda)\mathcal{B}) = \sum_{h=1}^{n} e(\mathbf{P}; x \to h)\frac{\lambda^{h-1}}{(h-1)!}\frac{(1-\lambda)^{n-h}}{(n-h)!}.$$

This can be interpreted as expressing $\mathrm{Vol}_{n-1}(\lambda\mathcal{A} + (1-\lambda)\mathcal{B})$ as a polynomial in λ, whose coefficients are given in terms of the numbers $e(\mathbf{P}; x \to h)$.

Now we can state the key results from convex geometry. For $\mathcal{A}$ and $\mathcal{B}$ convex bodies in $\mathbb{R}^m$, define their *joint dimension* to be the dimension of the affine hull of $\lambda\mathcal{A} + \mu\mathcal{B}$, for any $\lambda, \mu > 0$: this is well-defined.

Theorem 7.5.1 (Minkowski's Theorem) *Let $\mathcal{A}$ and $\mathcal{B}$ be two convex bodies in $\mathbb{R}^m$, and let d be their joint dimension. Then there are non-negative* *constants* $V(\mathcal{A}^k\mathcal{B}^{d-k})$*, for* $k = 0, \ldots, d$*, called the* mixed volumes *of* $\mathcal{A}$ *and* $\mathcal{B}$*, such that*

$$\mathrm{Vol}_d(\lambda\mathcal{A} + \mu\mathcal{B}) = \sum_{k=0}^{d} \binom{d}{k} V(\mathcal{A}^k\mathcal{B}^{d-k})\lambda^k\mu^{d-k}$$

for all non-negative constants λ*,* μ*.*

Setting $\mu = 0$, $\lambda = 1$, gives that $V(\mathcal{A}^d\mathcal{B}^0) = \mathrm{Vol}_d(\mathcal{A})$, and similarly $V(\mathcal{A}^0\mathcal{B}^d) = \mathrm{Vol}_d(\mathcal{B})$. The other mixed volumes don't generally have any such simple interpretation.

In our example, $d = n-1$, and $V(\mathcal{A}^k\mathcal{B}^{n-1-k}) = e(\mathbf{P}; x \to k+1)/(n-1)!$, for $k = 0, \ldots, n-1$.

Theorem 7.5.2 (Alexandrov-Fenchel Inequalities for Mixed Volumes) *For any pair* $\mathcal{A}, \mathcal{B}$ *of convex bodies with joint dimension* d*, the sequence* $V(\mathcal{A}^k\mathcal{B}^{d-k})$*,* $k = 0, \ldots, d$*, is log-concave.*

Applying this in our situation gives the result of Stanley.

Theorem 7.5.3 *For any* n*-element poset* $\mathbf{P} = (X, P)$*, and any* $x \in X$*, the sequence* $e(\mathbf{P}; x \to h)$*,* $h = 1, \ldots, n$*, is log-concave.*

Given that this result is true, one would expect there to be an elementary proof, perhaps giving an injection from $E(\mathbf{P}; x \to h+1) \times E(\mathbf{P}; x \to h-1)$ to $E(\mathbf{P}; x \to h) \times E(\mathbf{P}; x \to h)$, where $E(\mathbf{P}; x \to k)$ is the set of linear extensions $\prec$ of P with $h_\prec(x) = k$. But Stanley's proof is the only one known.

Minkowski's Theorem and the Alexandrov-Fenchel Inequalities can be generalized in the obvious way to apply to sums of more than two bodies, as we now state.

Theorem 7.5.4 *Let* $\mathcal{A}_1, \ldots, \mathcal{A}_r$ *be convex bodies in* $\mathbb{R}^m$*, and let* d *be the dimension of* $\mathcal{A}_1 + \cdots + \mathcal{A}_r$*. Then there is a set of non-negative constants* $V(\mathcal{A}_1^{k_1} \cdots \mathcal{A}_r^{k_r})$*, one for each sequence* $(k_1, \ldots, k_r)$ *of non-negative integers summing to* d*, such that*

$$\mathrm{Vol}_d(\lambda_1\mathcal{A}_1 + \cdots + \lambda_r\mathcal{A}_r) = \sum_{(k_1,\ldots,k_r)} \binom{d}{k_1, \ldots, k_r} V(\mathcal{A}_1^{k_1} \cdots \mathcal{A}_r^{k_r})\lambda_1^{k_1} \cdots \lambda_r^{k_r},$$

for any non-negative $\lambda_1, \ldots \lambda_r$*.*

Moreover, if $k_3 + \cdots + k_r = d - k$*, then the sequence* $V(\mathcal{A}_1^h\mathcal{A}_2^{k-h}\mathcal{A}_3^{k_3} \cdots \mathcal{A}_r^{k_r})$*,* $h = 0, \ldots, k$*, is log-concave.*

An application of this result is given in Exercise 18.

Just as for the lattice methods we considered earlier in the chapter, one cannot expect to be able to apply the Alexandrov-Fenchel Inequalities too often: there is no reason to suppose that a typical sequence arising in a combinatorial setting will be a sequence of mixed volumes. But when the technique works, it is extremely powerful.

Here is a result, and (sketch) proof, of Kahn and Saks [99], along extremely similar lines to Theorem 7.5.3.

Theorem 7.5.5 *Let $\mathbf{P} = (X, P)$ be an n-element poset, and let x and y be distinct elements of X. For $i = 1, 2, \ldots, n$, let $f(i)$ be the number of linear extensions $\prec$ of P in which $h_\prec(y) - h_\prec(x) = i$. The sequence $f(i)$, $i = 1, \ldots, n$, is log-concave.*

Proof We give just an outline of the proof, leaving the details to the reader (Exercise 19).

For $\lambda \in [0, 1]$, let $\mathcal{O}(\mathbf{P}; x, y; \lambda) = \{\mathbf{a} \in \mathcal{O}(\mathbf{P}) : a_y - a_x = \lambda\}$.

Let $\prec$ be a linear extension of P in which $h_\prec(y) - h_\prec(x) = i$. Then one can show that

$$\mathrm{Vol}_{n-1}(\mathcal{O}(\prec; x, y; \lambda)) = \frac{\lambda^{i-1}}{(i-1)!} \frac{(1-\lambda)^{n-i}}{(n-i)!}.$$

Therefore

$$\mathrm{Vol}_{n-1}(\mathcal{O}(\mathbf{P}; x, y; \lambda)) = \sum_{i=1}^{n} f(i) \frac{\lambda^{i-1}}{(i-1)!} \frac{(1-\lambda)^{n-i}}{(n-i)!}.$$

Also $\mathcal{O}(\mathbf{P}; x, y; \lambda) = \lambda \mathcal{O}(\mathbf{P}; x, y; 1) + (1-\lambda)\mathcal{O}(\mathbf{P}; x, y; 0)$. So $f(i)$ is equal to the mixed volume $V(\mathcal{O}(\mathbf{P}; x, y; 1)^{i-1} \mathcal{O}(\mathbf{P}; x, y; 1)^{n-i})/(n-1)!$, for $i = 1, \ldots, n$, and the result follows by the Alexandrov-Fenchel Inequalities.

(Notice that we have included the term $f(n)$, although it is always zero. This corresponds to the fact that the $(n-1)$-dimensional volume of $\mathcal{O}(\mathbf{P}; x, y; 1)$ is always equal to zero.) □

We conclude this chapter by turning back to order-preserving maps. Whatever intuition we might have for the sequence $e(\mathbf{P}; x \to h)$ being log-concave or unimodular, surely this applies equally if we deal with order-preserving maps instead of linear extensions? One might try to use Theorem 7.5.3 to prove the analogous result for order-preserving maps, or one might try to find a proof using mixed volumes. Neither of these approaches seem to work, but Daykin, Daykin and Paterson [99] found an ingenious elementary proof of the desired result.

Theorem 7.5.6 *For a poset* $\mathbf{P} = (X, P)$*, an element* $x \in X$*, a natural number* k*, and* $j \in [k]$*, let* $f(j) = f(\mathbf{P}; k; x \to j)$ *be the number of order-preserving maps* $\omega \in M_k(\mathbf{P})$ *with* $\omega(x) = j$*.*

The sequence $f(j)$*,* $j = 1, \dots, k$*, is log-concave.*

Proof For $\mathbf{P} = (X, P)$, x, k and j as above, let $F(j) = F(\mathbf{P}; k; x \to j)$ be the set of order-preserving maps $\omega \in M_k(\mathbf{P})$ with $\omega(x) = j$. We shall prove the result by giving an explicit injection Φ from $F(j-1) \times F(j+1)$ to $F(j) \times F(j)$, for $j = 2, \dots, n-1$.

Given any pair of order-preserving maps (ω_1, ω_2) in $M_k(\mathbf{P})$, and any subset Y of X, we say that the following pair (ω_3, ω_4) of maps is obtained by *exchanging* Y:

$$\text{for } z \notin Y, \begin{cases} \omega_3(z) = \omega_1(z) \\ \omega_4(z) = \omega_2(z); \end{cases} \qquad \text{for } y \in Y, \begin{cases} \omega_3(y) = \omega_2(y) - 1 \\ \omega_4(y) = \omega_1(y) + 1. \end{cases}$$

Note that the range of ω_3 may include 0, and that of ω_4 may include $k+1$. Note also that, if we take the new maps (ω_3, ω_4), and exchange the same set Y, then we recover (ω_1, ω_2).

Now suppose our initial maps ω_1, ω_2 are in $F(j-1)$ and $F(j+1)$ respectively. We shall choose a suitable set $Y = Y(\omega_1, \omega_2)$ containing x, and then set $\Phi(\omega_1, \omega_2)$ to be the pair (ω_3, ω_4) obtained by exchanging Y. We would like these two new maps both to be in $F(j) = F(\mathbf{P}; k; x \to j)$; choosing Y to include x does ensure that $\omega_3(x) = \omega_4(x) = j$.

How can ω_3 fail to be order-preserving? One possibility is that there are elements $y \in Y$, $z \notin Y$, with $y > z$, but $\omega_2(y) - 1 = \omega_3(y) < \omega_3(z) = \omega_1(z)$. Alternatively, there are $y \in Y$, $z \notin Y$, with $y < z$, but $\omega_2(y) - 1 > \omega_1(z)$. Similarly, if ω_4 fails to be order-preserving, then there are $y \in Y$, $z \notin Y$, such that either $y > z$ but $\omega_1(y) + 1 < \omega_2(z)$, or $y < z$ but $\omega_1(y) + 1 > \omega_2(z)$.

Accordingly, we say that y *forces* z (with respect to (ω_1, ω_2)) if any of the following four applies:

(i) $y > z$, $\omega_2(y) - 1 < \omega_1(z)$;
(ii) $y < z$, $\omega_2(y) - 1 > \omega_1(z)$;
(iii) $y > z$, $\omega_1(y) + 1 < \omega_2(z)$;
(iv) $y < z$, $\omega_1(y) + 1 > \omega_2(z)$.

We then define the set $Y = Y(\omega_1, \omega_2)$ to be the unique minimal subset of X containing x and closed under forcing. Equivalently, Y is the unique minimal set containing x whose exchange results in a pair of order-preserving maps.

We next claim that all elements $y \in Y$ satisfy $\omega_1(y) + 1 < \omega_2(y)$: in particular, this implies that the images of ω_3 and ω_4 are contained in $[k]$.

To prove the claim, we first notice that x satisfies the condition. Thus we need only prove that, if the condition holds for y, and y forces z, then it holds for z. If y and z are as in (i) above, then $\omega_2(y) - 1 < \omega_1(z) \le \omega_1(y)$, since ω_1 is order-preserving, so y does not satisfy the condition. If y and z are as in (ii) above, then $\omega_2(z) - 1 \ge \omega_2(y) - 1 > \omega_1(z)$, so z does satisfy the condition. We argue similarly if y and z are as in (iii) or (iv) above.

We have now shown that ω_3 and ω_4 are both in $F(j)$, so Φ is a map from $F(j-1) \times F(j+1) \to F(j) \times F(j)$. To complete the proof, we need to show that Φ is an injection, i.e., (ω_1, ω_2) can be recovered from (ω_3, ω_4).

Suppose we are given (ω_3, ω_4); as before, there is a unique minimal set Z, containing x, whose exchange gives a pair of order-preserving maps. We know that exchanging Y gives us back (ω_1, ω_2), so $Z \subseteq Y$. We claim that $Z = Y$: to prove this, it is enough to show that, whenever $y \in Y$ forces z with respect to (ω_1, ω_2), then also y forces z with respect to (ω_3, ω_4).

This last claim is just a matter of rewriting: if $y \in Y$ forces z with respect to (ω_1, ω_2), then both elements are in Y, and one of the following four conditions holds:

(i) $y > z$, $\omega_3(y) < \omega_4(z) - 1$;
(ii) $y < z$, $\omega_3(y) > \omega_4(z) - 1$;
(iii) $y > z$, $\omega_4(y) < \omega_3(z) + 1$;
(iv) $y < z$, $\omega_4(y) > \omega_3(z) + 1$.

In each of these cases, y does indeed force z with respect to (ω_3, ω_4). □

See Exercise 20 for an example to illustrate this proof.

Exercises

7.1 Show that, for any subsets A and B of a set S, the following are equivalent. (Here $\overline{C} = S \setminus C$, for $C \subseteq S$.)

$$\begin{aligned} |A|\,|B| &\le |A \cap B|\,|S|, \\ |A|\,|\overline{B}| &\ge |A \cap \overline{B}|\,|S|, \\ |A \cap \overline{B}|\,|B \cap \overline{A}| &\le |A \cap B||\overline{A} \cap \overline{B}|. \end{aligned}$$

7.2 In the FKG Inequality, is the condition that f and g be functions to the *non-negative* reals essential? What about the condition that μ is a function to the *non-negative* reals?

7.3 Prove that, if f, g and h are functions from $\mathbf{n}$ to $\mathbb{R}^+$ with f and g non-decreasing, then

$$\left(\sum_{i=0}^{n-1} f(i)h(i)\right)\left(\sum_{i=0}^{n-1} g(i)h(i)\right) \le \left(\sum_{i=0}^{n-1} f(i)g(i)h(i)\right)\left(\sum_{i=0}^{n-1} h(i)\right).$$

This is *Chebychev's Inequality.*

7.4 Suppose that $\mathcal{A}$ and $\mathcal{B}$ are two subsets of $\mathbf{2}^n$ such that, for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$, A and B are incomparable in the subset order. Show that $|\mathcal{A}|\,|\mathcal{B}| \leq 2^{2n-4}$. (*Hint: use (7.3).*)

7.5 Suppose that $\mathcal{A}$ is a subset of $\mathbf{2}^n$ such that, for every pair A, B of sets in $\mathcal{A}$, $A \cap B \neq \emptyset$, and $A \cup B \neq [n]$. Show that $|\mathcal{A}| \leq 2^{n-2}$. Show also that this inequality is best possible.

7.6 Show that, for any non-distributive lattice $\mathbf{L}$, there are subsets A and B of $\mathbf{L}$ such that $|A|\,|B| > |A \vee B|\,|A \wedge B|$.

7.7 Show that, for any up-sets U, V of a lattice $\mathbf{L}$, $U \vee V = U \cap V$.

7.8 Explain why, for any poset $\mathbf{P}$, $M_2(\mathbf{P})$ is in 1-1 correspondence with the down-set lattice $\mathcal{D}(\mathbf{P})$.

7.9 Let $\mathbf{P} = (X, P)$ be a poset covered by two disjoint chains Y and Z. Show that it need not be the case that, in $M_2(\mathbf{P})$, events $\omega(y_1) \prec \omega(z_1)$ and $\omega(y_2) \prec \omega(z_2)$ are non-negatively correlated whenever $y_1, y_2 \in Y$ and $z_1, z_2 \in Z$. (*Hint: there is an example with* $|X| = 4$.)

7.10 How many order-preserving maps are there from $[n]$ to $[k]$?

7.11 Let $\mathbf{P} = (X, P)$ be a finite poset. Show that, as $k \to \infty$,

$$\Pr{}_k(\omega(x) = \omega(y) \text{ for some } x, y \in X) \to 0.$$

7.12 For the lattice structure $\mathbf{L}$ on $M_k(\mathbf{P})$ defined in the proof of Theorem 7.3.3, which are the join-irreducible elements? Describe the poset $\mathbf{Q}$ such that $\mathbf{L}$ is isomorphic to $\mathbf{2}^{\mathbf{Q}}$. What is the dimension of $\mathbf{L}$?

7.13 Give $M_k(\mathbf{P})$ the lattice structure arising from the embedding Φ of $M_k(\mathbf{P})$ into $\mathcal{H}_k$, as in Theorem 7.3.5. For $-k+2 \leq \ell \leq k-1$, let $\mathcal{H}_k^\ell$ be the sublattice of $\mathcal{H}_k$ consisting of those vectors whose second co-ordinate is ℓ or $\ell - 1$. By considering the pre-image of $\mathcal{H}_k^\ell$ under Φ, show that the function

$$\Pr{}_k(\omega(x) \geq m \mid \omega(x) - \omega(y) = \ell)$$

is non-decreasing in ℓ, for each fixed k.

7.14 Suppose that x, y, z are elements of a finite poset $\mathbf{P}$. Show that

$$\Pr{}_k(\omega(x) \preceq \omega(y) \mid \omega(y) = \ell, \omega(z) = m)$$

is non-decreasing in ℓ, for fixed k and m.

7.15 Is it true that, for every poset $\mathbf{P} = (X, P)$ and every four elements a, b, c, d, the events $a \prec c \prec b$ and $a \prec d \prec b$ are non-negatively correlated?

7.16 Suppose that the sequence $f(j)$ is log-concave. Show that $f(i)f(i+k+\ell) \le f(i+k)f(i+\ell)$ for any i, and any $k, \ell \ge 0$.
(*Sometimes this is given as the definition of log-concavity: this exercise shows that it is enough to verify the condition for* $k = \ell = 1$.)

7.17 Use Theorem 7.5.3 to show that, for fixed m, the sequence of binomial coefficients $\binom{m}{h}$, $h = 0, \dots, m$, is log-concave.
(*Unsurprisingly, this is easy to prove directly.*)

7.18 Let x and y be elements of an n-element poset $\mathbf{P} = (X, P)$, with $x < y$, and let m be an element of $[n]$ such that $e(\mathbf{P}; y \to m) > 0$. Let $f(h)$ be the number of linear extensions $\prec$ in which $h_\prec(x) = h$ and $h_\prec(y) = m$.

Use Theorem 7.5.4 to show that the sequence $f(h)$, $h = 1, \dots, m-1$, is log-concave.

7.19 Fill in the details in the proof of Theorem 7.5.5.

7.20 Let $\mathbf{P} = (X, P)$ be the three-element poset on $\{x, y, z\}$ whose only relations are $x < y$ and $z < y$. Explore how the proof of Theorem 7.5.6 applies to this example.

State explicitly which pairs $(\omega_1, \omega_2) \in F(j-1) \times F(j+1)$ have Y equal to (a) $\{x\}$, (b) $\{x, y\}$, (c) $\{x, y, z\}$. Verify that the images of Φ in these three cases are disjoint.

Give an example of a pair $(\omega, \omega') \in F(j) \times F(j)$ that is not in the image of Φ.

7.6 Notes and References

The result in Exercise 4 is due to Seymour [99]. Many sources in the literature credit Seymour with proving (7.2), although this follows instantly from Kleitman's result, and Seymour's paper, organized as an application of Kleitman's Theorem, does not draw a distinction between the two forms.

Theorem 7.3.3 appears in Shepp [99], with the proof given here. Brightwell [99] showed that one has strict positive correlation here unless all of y_1, y_2, z_1 and z_2 are in different components of the comparability graph of $\mathbf{P}$. The following question is also answered in [99]: for which posets $\mathbf{P} = (X, P)$ and 4-tuples of elements (x, y, u, v) is it the case that, for every poset $\mathbf{Q} = (X, Q)$ with $Q \subseteq P$, the events $x \prec y$ and $u \prec v$ are non-negatively correlated? (For instance, by Theorem 7.3.3, this is true whenever neither x nor y shares a component of $\mathbf{P}$ with u or v.) The answer is that this is the case exactly when, in the comparability graph of $\mathbf{P}$, every path from x to v passes through either y or u, and every path from y to u passes through either x or v. Exercises 13 and 14 are lemmas from [99].

Theorem 7.3.4 is due to Fishburn [99], who needed it as a lemma in his proof of the strict xyz Inequality; the proof here, via the theorem of Shepp, was given by Brightwell [99].

Theorem 7.4.1 was proved by Graham, Yao and Yao [99]. Much shorter proofs, using the FKG Inequality or its relatives, were given by Kleitman and Shearer [99] and by Shepp [99]. Brightwell [99] gave a proof via an explicit injection, and also showed that, under the conditions of the theorem, events $y_1 \prec z_1$ and $y_2 \prec z_2$ are strictly positively correlated except in trivial cases.

There is a conjectured common generalization of Theorem 7.4.1 and Theorem 7.3.3: if $\mathbf{P}$ is obtained as a lexicographic sum by taking a union of two chains, Y and Z, and substituting posets Y_i and Z_j for each element $y_i \in Y$ and $z_j \in Z$, and S and T are subsets of $(\cup Y_i) \times (\cup Z_j)$, then G_S and G_T are non-negatively correlated. This was first conjectured by Graham [99] in a different form, shown to be equivalent to this one by Daykin and Daykin [99]. The fact that this is resistant to various attempts to prove it based on the correlation inequalities introduced in this chapter perhaps reveals that we are still some way from a complete understanding of the phenomenon of correlation in posets.

The xyz Inequality was originally conjectured by Rival and Sands [99], and proved a few months later by Shepp. The proofs of the xyz Inequality here are almost exactly as they appear in the papers of Shepp [99] and Fishburn [99].

Theorem 7.5.1 was proved by Minkowski [99] in 1910, and the Alexandrov-Fenchel Inequalities were proved independently by Alexandrov [99] and Fenchel [99] in 1936. By now, there are a number of proofs of the Alexandrov-Fenchel Inequalities, but none that is suitable for description here. A recent book covering this area is that of Schneider [99]. The proofs of Theorem 7.5.3 and Theorem 7.5.5 we give here are as in the original papers of Stanley [99] and Kahn and Saks [99] respectively.

As mentioned in the course of the chapter, it would be good to have more methods for proving correlation or log-concavity results that don't require the construction of a distributive lattice or a sequence of mixed volumes. As a first step in this direction, Brightwell and Trotter [99] give elementary proofs of Shepp's result, Theorem 7.3.3, and of Fishburn's strong form of the xyz Inequality. These proofs are not shorter or simpler than the ones presented here, but they do not use any version of the Ahlswede-Daykin Theorem, and it is hoped that they might be generalized to settings where it is not possible to construct a convenient auxiliary lattice. As yet, however, there are no successes to report.