Symmetric polynomials and symmetric mean inequalities

Karl Mahlburg *
Department of Mathematics
Louisiana State University
Baton Rouge, LA 70803
mahlburg@math.lsu.edu

Clifford Smyth
Department of Mathematics and Statistics
University of North Carolina Greensboro
Greensboro, NC 27402
cdsmyth@uncg.edu

Submitted: XX; Accepted: XX; Published: XX
Mathematics Subject Classifications: 05E05, 26E60, 60C05

Abstract

We prove generalized arithmetic-geometric mean inequalities for quasi-means arising from symmetric polynomials. The inequalities are satisfied by all positive, homogeneous symmetric polynomials, as well as a certain family of non-homogeneous polynomials; this family allows us to prove the following combinatorial result for marked square grids.

Suppose that the cells of a $n \times n$ checkerboard are each independently filled or empty, where the probability that a cell is filled depends only on its column. We prove that for any $0 \leq \ell \leq n$, the probability that each column has at most $\ell$ filled sites is less than or equal to the probability that each row has at most $\ell$ filled sites.

Keywords: symmetric means; symmetric polynomials; arithmetic-geometric mean inequality

MSC: 60C05, 05E05, 26E60, 05C80

1 Introduction

Let $n$ be a positive integer. We define an $n$-variable orthant function to be a continuous function $F : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ such that $F(x) = F(x_1, \ldots, x_n)$ is monotonically increasing in each $x_i$ and that also has a strictly increasing diagonal restriction, $f_F(y) = F(y, \ldots, y)$.

*Partially supported by an NSF Postdoctoral Fellowship administered by the Mathematical Sciences Research Institute through its core grant DMS-0441170 and by NSF Grant DMS-1201435..
Given an $n$-variable orthant function $F$, we define the following $n$-variable orthant functions associated with $F$: the quasi-arithmetic mean, the quasi-geometric mean, and the quasi-mean; respectively, they are

\[ A_F(x) := f_F^{-1}\left(\frac{f_F(x_1) + \cdots + f_F(x_n)}{n}\right), \]

\[ G_F(x) := f_F^{-1}\left(\prod_{j=1}^{n} f_F(x_j)^{\frac{1}{j}}\right), \]

\[ M_F(x) := f_F^{-1}(F(x)). \]

Note that these means have been studied classically (see [2], Chapter III).

Some care is needed to verify that these definitions are well-defined. One must note that since $f_F(y)$ is strictly increasing and continuous, its range $R := f_F(\mathbb{R}_{\geq 0})$ is of the form $R = [f_F(0), M]$ or $[f_F(0), +\infty)$, according to the value $M = \lim_{y \to +\infty} f_F(y)$. Furthermore, $f_F$ is a bijection and $f_F^{-1}$ is strictly increasing and continuous. Since $R$ is closed under taking arithmetic and geometric means of its elements, $A_F(x)$ and $G_F(x)$ are well-defined. Since $F$ is monotonically increasing in each variable, it also satisfies

\[ f_F(x) \leq F(x) \leq f_F(\overline{x}) \]

where $\underline{x} := \min\{x_i : 1 \leq i \leq n\}$ and $\overline{x} := \max\{x_i : 1 \leq i \leq n\}$. This implies that $F(\mathbb{R}_{\geq 0}) = R$ and so $M_F(x)$ is well-defined also.

If $M = A_F, G_F,$ or $M_F$, we therefore have that $M$ is a quasi-mean. In particular, $M$ satisfies the following usual properties of a mean: it is continuous and monotonically increasing in each variable, $\underline{x} \leq M(x) \leq \overline{x}$ for all $x \in \mathbb{R}_{\geq 0}$, and $M(y, \ldots, y) = y$ for all $y \geq 0$. Note that the $f_F^{-1}$ in the definition of $M$ ensures the final “identity” property, $M(y, \ldots, y) = y$. However, $M$ is not necessarily linearly homogeneous, i.e. we do not necessarily have $M(\lambda x) = \lambda M(x)$ for all $\lambda \geq 0$ and $x \in \mathbb{R}_{\geq 0}$. The function $F(x) = (1 + x_1)(1 + x_2)$ provides a simple counterexample for each type of $M$.

Since $f_F$ and hence $f_F^{-1}$ are strictly increasing, $A_F$ and $G_F$ are strictly increasing in each variable. An arithmetic-geometric mean inequality between $G_F$ and $A_F$ also easily follows: for all $x \in \mathbb{R}_{\geq 0}^n$,

\[ G_F(x) \leq A_F(x), \]

with equality if and only if all $x_i$ are equal. This is also Theorem 85 of [2], with $\psi = \log f_F$, $\chi = f_F$, and $q \equiv 1/n$. Note that the classical arithmetic-geometric mean inequality can be recovered from this by setting $F(x) = (x_1 \cdots x_n)^{1/n}$.

In this paper we study functions whose quasi-means provide a refinement of the preceding arithmetic-geometric mean inequality. Namely, we are interested in $S$, which we define to be the set of all orthant functions $F$ for which

\[ G_F(x) \leq M_F(x) \leq A_F(x), \quad \text{for all } x \in \mathbb{R}_{\geq 0}^n. \]

Proposition 1.1. The set $S$ satisfies the following properties.
1. If \( F(x_1, \ldots, x_n) \) is a homogeneous symmetric polynomial with positive coefficients, then \( F \in S \).

2. If \( F_i(x_1, \ldots, x_n) \in S \) for \( i = 1, 2 \), then \( F_1 \cdot F_2 \in S \).

Remark 1.2. The set \( S \) is not closed under addition. For example, \( F(x) = 1 + x_1 x_2 \) is the sum of two homogeneous functions but \( G_F \leq M_F \) fails to hold when \( x_1 \neq x_2 \).

For integers \( \ell \) and \( n \) with \( 0 \leq \ell \leq n \) we define

\[
\eta_\ell(x) := \sum_{j=0}^{\ell} \sum_{\substack{I \subseteq [n] \mid |I| = j}} x_I, \tag{1.3}
\]

where we have adopted the common notation \( x_I := \prod_{i \in I} x_i \), along with the convention \( x_{\emptyset} = 1 \). Note that \( \eta_\ell \) is an orthant function if and only if \( \ell \geq 1 \). We denote the diagonal restriction of \( \eta_\ell \) by

\[
\mu_\ell(y) := \eta_\ell(y, \ldots, y) = \sum_{j=0}^{\ell} \binom{n}{j} y^j. \tag{1.4}
\]

In Section 3 we will prove the following result, which states that \( \eta_\ell \in S \) for \( 1 \leq \ell \leq n \).

**Theorem 1.3.** If \( \ell \) and \( n \) are integers with \( 0 \leq \ell \leq n \) and \( n \geq 1 \), then

\[
\left( \prod_{i=1}^{n} \mu_\ell(x_i) \right)^{1/n} \leq \eta_\ell(x) \leq \frac{1}{n} \sum_{i=1}^{n} \mu_\ell(x_i), \text{ for all } x \in \mathbb{R}_{\geq 0}^n.
\]

*Equality holds throughout if and only if \( \ell = 0, \ell = n \), or \( x_1 = \cdots = x_n \).*

These quasi-mean inequalities have an appealing application to combinatorial probability. Let \( \{X_{ij}, 1 \leq i, j \leq n\} \) be a collection of independent Bernoulli random variables with probability \( p_j \); in other words, \( P(X_{ij} = 1) = p_j \) and \( P(X_{ij} = 0) = 1 - p_j \). Using these, we further define the random variables

\[
C_j := \sum_{i=1}^{n} X_{ij} \quad \text{for } 1 \leq j \leq n,
\]

\[
R_i := \sum_{j=1}^{n} X_{ij} \quad \text{for } 1 \leq i \leq n.
\]

Using Theorem 1.3, we prove bounds relating the distributions of the \( C_j \)'s and \( R_i \)'s.

**Theorem 1.4.** Suppose that \( p_1, \ldots, p_n \in [0, 1] \). If \( \ell \) is an integer with \( 0 \leq \ell \leq n \), then

\[
P \left( \max_{1 \leq j \leq n} \{C_j\} \leq \ell \right) \leq P \left( \max_{1 \leq i \leq n} \{R_i\} \leq \ell \right)
\]
Theorem 1.4 has two immediate combinatorial reformulations: one to marked square grids and another to random bipartite graphs. First, suppose that markers are independently placed in the squares of an \( n \) by \( n \) grid such that the probability that a marker is placed in a square depends only on that square’s column. Then for all integers \( \ell \) with \( 0 \leq \ell \leq n \), the probability that every column has at most \( \ell \) markers is less than or equal to the probability that each row has at most \( \ell \) markers.

Alternatively, suppose that \( G \) is a finite complete bipartite graph with bipartition \((A,B)\), where \(|A|=|B|=n\). Let \( H \) be a random subgraph of \( G \) where each edge \( e \) of \( G \) is independently selected to belong to \( H \) with a probability that depends only on the left vertex \( e \cap A \). If \( \ell \geq 0 \), then

\[
P \left( \max_{a \in A} \{ d_H(a) \} \leq \ell \right) \leq P \left( \max_{b \in B} \{ d_H(b) \} \leq \ell \right).
\]

Remark 1.5. In the grid formulation, both events in the inequality require that at most \( n\ell \) squares be occupied. Similarly, both events in the bipartite graph formulation require that there be at most \( n\ell \) edges.

The remainder of the paper is as follows. In Section 2 we study the basic properties of homogeneous symmetric polynomials and the functions in \( S \), and also prove Proposition 1.1. In Section 3 we use Lagrange multipliers and polynomial inequalities to prove Theorem 1.3. We conclude in Section 4, where we describe the relationship between the quasi-mean inequalities and combinatorial probability, and prove Theorem 1.4.

## 2 Properties of symmetric polynomials and \( S \)

As found in Chapter 7 of [4], one can define various homogeneous (graded) bases for the ring of symmetric polynomials on \( n \) variables; we will make explicit use of the *elementary symmetric polynomials* \( \{ e_j(x) \mid 0 \leq j \leq n \} \), where the degree \( j \) polynomial is defined as

\[
e_j(x) := \sum \{ x_I : I \subseteq [n], |I| = j \}.
\]

We will also use the *monomial symmetric polynomials*, which are defined as follows.

For a positive integer \( n \), let \( \Delta(n) := \{ \lambda \in \mathbb{R}_{\geq 0}^n : \lambda_1 \geq \cdots \geq \lambda_n \} \), and define the *weight* of such a vector \( \lambda \) as \(|\lambda| := \lambda(1) + \cdots + \lambda(n)\). If \( \lambda \in \Delta(n) \), the monomial symmetric polynomial associated to \( \lambda \) is defined as

\[
M_\lambda(x) := \frac{1}{n!} \sum_{\sigma \in S_n} x_{\sigma(1)}^{\lambda(1)} \cdots x_{\sigma(n)}^{\lambda(n)}.
\tag{2.1}
\]

Note that this is homogeneous of degree \(|\lambda|\).

We now recall various inequalities between symmetric polynomials. Suppose that \( \lambda_i = (\lambda_i(1), \ldots, \lambda_i(n)) \in \Delta(n) \) for \( i \in \{1,2\} \). We say that \( \lambda_1 \) *majorizes* \( \lambda_2 \) if and only if \(|\lambda_1| = |\lambda_2|\), and

\[
\lambda_1(1) + \cdots + \lambda_1(j) \geq \lambda_2(1) + \cdots + \lambda_2(j)
\]

for all \( 1 \leq j < n \). In this case, we write \( \lambda_1 \succeq \lambda_2 \). Muirhead’s inequalities can be concisely stated as follows.
Theorem (Muirhead [3]). Suppose that $\lambda_1, \lambda_2 \in \Delta(n)$. The inequality
\[ M_{\lambda_1}(x) \geq M_{\lambda_2}(x) \] (2.2)
is satisfied for all $x \in \mathbb{R}^n_{\geq 0}$ if and only if $\lambda_1 \succeq \lambda_2$.

Note that [3] only contains the case where both $\lambda_i$ have integral parts, while [2] (Theorem 45, p. 45) contains the general result above. We also recall the following elementary result from the general theory of series inequalities; see [2] (Theorem 368, p. 261).

Theorem (Rearrangement Inequality). Suppose that $a_1 \geq \cdots \geq a_n \geq 0$ and $b_1 \geq \cdots \geq b_n \geq 0$. If $\sigma \in S_n$ is a permutation, then
\[ \sum_{i=1}^{n} a_i b_i \geq \sum_{i=1}^{n} a_i b_{\sigma(i)}. \] (2.3)

Our final preliminary observations address the inequalities in (1.2) individually. Let $\mathcal{L}$ denote the set of orthant functions $F$ that satisfy the left inequality:
\[ \mathcal{G}_F(x) \leq M_F(x), \quad \text{for all } x \in \mathbb{R}^n_{\geq 0}, \]
and let $\mathcal{R}$ denote the set of orthant functions that satisfy the right inequality:
\[ M_F(x) \leq A_F(x), \quad \text{for all } x \in \mathbb{R}^n_{\geq 0}. \]

Clearly we have $\mathcal{S} = \mathcal{L} \cap \mathcal{R}$. The following properties follow immediately from the definitions of the quasi-means.

Proposition 2.1. The classes $\mathcal{L}$ and $\mathcal{R}$ have the following closure properties:

1. If $F_i \in \mathcal{L}$ for $i = 1, 2$, then $F_1^a \cdot F_2^b \in \mathcal{L}$ for all $a, b \geq 0$.
2. If $F_i \in \mathcal{R}$ for $i = 1, 2$, then $aF_1 + bF_2 \in \mathcal{R}$ for all $a, b \geq 0$.

The preceding facts now allow us to prove our first result about $\mathcal{S}$.

Proof of Proposition 1.1. We first prove statement 1. Suppose $F$ is a homogeneous symmetric polynomial with positive coefficients and total degree $w$. It can be written as
\[ F(x) = \sum_{i=1}^{k} a_i M_{\lambda_i}(x), \] (2.4)
where $a_i > 0$, $\lambda_i \in \Delta(n)$ and $|\lambda_i| = w$ for all $i$. Writing $A := \sum_{i=1}^{k} a_i > 0$, Muirhead’s theorem (2.2) then implies that
\[ A \cdot M_{(w/n, \ldots, w/n)}(x) \leq F(x) \leq A \cdot M_{(w, 0, \ldots, 0)}(x), \quad \text{for all } x \in \mathbb{R}^n_{\geq 0}. \] (2.5)
We finish the proof by showing that this is equivalent to the statement $F \in S$. Since $f_F(y) = F(y, \ldots, y) = Ay^w$, we find that the leftmost expression in (2.5) is the same as

$$A \cdot \frac{1}{n!} \cdot (x_1 \cdots x_n)^{w/n} \cdot n! = \left( \prod_{i=1}^{n} f_F(x_i) \right)^{1/n}.$$

Similarly, the rightmost expression in (2.5) equals

$$A \cdot \frac{1}{n!} \left( x_1^w + \cdots + x_n^w \right) \cdot (n-1)! = \frac{1}{n} \sum_{i=1}^{n} f_F(x_i).$$

This completes the first part of the proof.

We now turn to statement 2. Throughout this part of the proof we write $f_i$ as shorthand for $f_{F_i}$. By part 1 of Proposition 2.1 we need only show that if $F_i \in S$ for $i = 1, 2$, then $F_1 \cdot F_2 \in R$. This is equivalent to showing that

$$(F_1 F_2)(x) \leq \frac{1}{n} \sum_{i=1}^{n} f_{F_1 F_2}(x), \quad (2.6)$$

and we can immediately rewrite $f_{F_1 F_2} = f_1 \cdot f_2$. Using the assumption that $F_i \in R$, we find that the left side of (2.6) satisfies

$$(F_1 F_2)(x) \leq \left( \frac{1}{n} \sum_{i=1}^{n} f_1(x_i) \right) \left( \frac{1}{n} \sum_{i=1}^{n} f_2(x_i) \right) = \frac{1}{n^2} \sum_{i_1, i_2=1}^{n} f_1(x_{i_1}) f_2(x_{i_2}). \quad (2.7)$$

Define the one-step shift cyclical permutation by $\sigma(i) := i + 1$ for $1 \leq i \leq n - 1$, and $\sigma(n) := 1$. Reordering the $x_i$ if necessary so that $x_1 \geq \cdots \geq x_n \geq 0$, we then further rewrite the sum from (2.7) as

$$\frac{1}{n^2} \sum_{i_1, i_2=1}^{n} f_1(x_{i_1}) f_2(x_{i_2}) = \frac{1}{n^2} \sum_{j=0}^{n-1} \sum_{i=1}^{n} f_1(x_i) f_2(x_{\sigma^j(i)}).$$

The Rearrangement Inequality, (2.3), now implies that the largest term in the outer summation occurs when $j = 0$, so

$$(F_1 F_2)(x) \leq \frac{1}{n} \sum_{i=1}^{n} (f_1 f_2)(x_i), \quad (2.8)$$

which verifies (2.6). \qed
3 Proof of Theorem 1.3

3.1 Overview

In this section we prove Theorem 1.3, devoting most of our effort to the left inequality. In particular, we consider the level sets of $\eta_\ell(x)$ and then apply the technique of Lagrange multipliers in order to determine the extremal behavior of

$$U(x) := \prod_{j=1}^{n} \mu_\ell(x_j)^{\frac{1}{n}}.$$

For each $R$ in the range of $\eta_\ell$ on $\mathbb{R}_{\geq 0}$, define the surface $\Omega(R) = \{x \in \mathbb{R}_{\geq 0}^n : \eta_\ell(x) = R\}$. Given $d > 0$ and an integer $k$ with $1 \leq k \leq n$, we define $c_k(d) = d(e_1 + \cdots + e_k) \in \mathbb{R}^n$, where $e_i$ is the $i$-th standard basis vector. We refer to a point in $\mathbb{R}_{\geq 0}^n$ as a $k$-diagonal point if it is any coordinate permutation of a point of the form $c_k(d)$ for some $d > 0$.

Lemma 3.1. If $1 \leq \ell \leq n-1$, $1 < R < \infty$, and $z$ is a maxima of $U(x)$ on $\Omega(R)$ then $z$ is a $k$-diagonal point for some $k$ with $1 \leq k \leq n$.

We will prove Lemma 3.1 in Section 3.2 using the method of Lagrange multipliers (see Theorem 3.3).

By the symmetry of $f$ and $\eta_\ell$, if we restrict our attention to a single point $x \in \mathbb{R}_{\geq 0}^n$, we may assume that there is some $0 \leq k \leq n$ such that $x_1, \ldots, x_k > 0$ and $x_{k+1} = \cdots = x_n = 0$. With this in mind, we generalize the functions defined in (1.3) and (1.4) by setting

$$\eta_{\ell,k}(x) := \eta_\ell(x_1, \ldots, x_k, 0, \ldots, 0),$$

$$\mu_{\ell,k}(y) := \eta_\ell(y, \ldots, y, 0, \ldots, 0).$$

These are related to our earlier definitions by $\eta_\ell = \eta_{\ell,n}$ and $\mu_\ell = \mu_{\ell,n}$, and we also have the further relations

$$\eta_{\ell,k}(x) = \sum_{I \subseteq [k] \atop |I| \leq \ell} x_I, \quad \mu_{\ell,k}(y) = \eta_{\ell,k}(y, \ldots, y) = \sum_{j=0}^{\ell} \binom{k}{j} x^j,$$

and finally, $\eta_{\ell,k}(x) = \eta_{k,k}(x) = \prod_{i=1}^{k} (1 + x_i)$ if $\ell \geq k$.

Lemma 3.2. If $0 \leq \ell \leq n$, $1 \leq k \leq n$, and $y \geq 0$, then

$$\mu_{\ell,n}^k(y) \leq \mu_{\ell,k}(y).$$

The inequality is tight if and only if $\ell = 0$, $\ell = n$, $y = 0$ or $k = n$.

We will prove Lemmas 3.1 and 3.2 in Sections 3.2 and 3.3, respectively. We now show how they imply Theorem 1.3.
Proof of Theorem 1.3. If \( \ell = 0 \) or \( \mathbf{x} = 0 \), then all terms in the two inequalities are equal to 1. If \( x_1 = \cdots = x_n = y \) then all terms are identically equal to \( \mu_{\ell}(y) \). If \( \ell = n \), the left inequality is an identity and the right inequality is an application of the arithmetic-geometric mean inequality. In general, the right inequality follows from Proposition 1.1 part 1 and Proposition 2.1 part 2. So it suffices to prove the left inequality in the case where \( \mathbf{x} \in \mathbb{R}^n_{\geq 0}, \mathbf{x} \neq 0 \) and \( 1 \leq \ell \leq n-1 \).

Since \( \eta_{\ell} \) is strictly increasing in each variable, a non-zero \( \mathbf{x} \) is contained in the surface \( \Omega(R) \) with \( R = R(\mathbf{x}) = \eta_{\ell}(\mathbf{x}) > \eta_{\ell}(0) = 1 \). Since \( \eta_{\ell}(\mathbf{x}) \geq 1 + x_i \) for all \( 1 \leq i \leq n \), \( \Omega(R) \) is bounded. Since \( \eta_{\ell}(\mathbf{x}) \) is continuous, \( \Omega(R) \) is also closed, and hence compact.

This means that there exists at least one point \( \mathbf{z} \in \Omega(R) \) at which \( \mathbf{f} \) takes its maximum value on \( \Omega(R) \). Since \( R > 1, \mathbf{z} \neq 0 \) and, by Lemma 3.1, \( \mathbf{z} = \mathbf{c}_k(d) \), for some \( 1 \leq k \leq n \) and \( d > 0 \). By Lemma 3.2, if \( k < n \), then

\[
\mathbf{f}(\mathbf{z}) = (\mu_{\ell,n}(d))^{\frac{k}{n}} < \mu_{\ell,k}(d) = \eta_{\ell}(\mathbf{z}) = R.
\]

However, if \( k = n \), we have \( \mathbf{f}(\mathbf{z}) = \mu_{\ell}(d) = \eta_{\ell}(\mathbf{z}) = R \). 

\( \square \)

3.2 Lagrange multipliers and maxima

In this section we prove Lemma 3.1 using Proposition 3.3.1 of [1], p.284. We restate this result as Theorem 3.3, a form more suitable to our purposes.

**Theorem 3.3** (The method of Lagrange multipliers with inequality constraints.). Let \( f, h_1, \ldots, h_m, g_1, \ldots, g_r : \mathbb{R}^n \to \mathbb{R} \) be continuously differentiable functions. Suppose \( \mathbf{z} \) is a point at which \( \mathbf{f}(\mathbf{x}) \) has a local maximum over \( \Omega = \{ \mathbf{x} : h_1(\mathbf{x}) = \cdots = h_m(\mathbf{x}) = 0, g_1(\mathbf{x}), \ldots, g_r(\mathbf{x}) \geq 0 \} \). Suppose also that \( \mathbf{z} \) is regular, i.e. \( \{ \nabla h_1(\mathbf{z}), \ldots, \nabla h_m(\mathbf{z}) \} \cup \{ \nabla g_i(\mathbf{z}) : i \in A(\mathbf{z}) \} \) is a linearly independent set where \( A(\mathbf{x}) := \{ 1 \leq j \leq r : g_j(\mathbf{x}) = 0 \} \). Then there exist unique Lagrange multiplier vectors \( \mathbf{\lambda}' \in \mathbb{R}^m \) and \( \mathbf{\rho}' \in \mathbb{R}^r_{\geq 0} \), such that

\[
\frac{\partial \mathbf{L}}{\partial x_i}(\mathbf{z}, \mathbf{\lambda}', \mathbf{\rho}') = 0, \quad 1 \leq i \leq n,
\]

\[
\rho_j' = 0, \quad j \notin A(z)
\]

where \( \mathbf{L}(\mathbf{x}, \mathbf{\lambda}, \mathbf{\rho}) := \mathbf{f}(\mathbf{x}) + \sum_{i=1}^m \lambda_i h_i(\mathbf{x}) + \sum_{j=1}^r \rho_j g_j(\mathbf{x}) \).

**Proof of Lemma 3.1.** Theorem 3.3 applies to the present setting with \( \Omega(R) \) as the set \( \Omega \), constraint function \( h(\mathbf{x}) := \eta_{\ell}(\mathbf{x}) - R \), and inequality constraints \( g_i(\mathbf{x}) := x_i \) for \( 1 \leq i \leq n \). The Lagrangian function is then

\[
\mathbf{L}(\mathbf{x}, \lambda, \mathbf{\rho}) := \mathbf{U}(\mathbf{x}) + \lambda(\eta_{\ell}(\mathbf{x}) - R) + \sum_{i=1}^n \rho_i x_i.
\]

Let \( \mathbf{z} \in \Omega(R) \) be a point at which \( \mathbf{U}(\mathbf{x}) \) attains its maximum value over \( \Omega(R) \). Since \( R > 1 \), \( \mathbf{z} \) has at least one non-zero coordinate. By symmetry we may assume \( z_1, z_2, \ldots, z_k > 0 \) and \( z_{k+1} = \cdots = z_n = 0 \) for some \( k \geq 1 \). Since \( \nabla h(\mathbf{z}) \) trivially has
positive coordinates, $\nabla h(z)$ together with $\nabla g_i(z) = e_i$, for $k + 1 \leq i \leq n$, form a linearly independent set. Thus $z$ is regular, and the conditions of Theorem 3.3 are met.

The theorem statement now implies that there is a constant $\lambda'$ such that

$$\frac{1}{n} \frac{\mu'_{\ell,n}(z_i)}{\mu_{\ell,n}(z_i)} U(z) + \lambda \frac{\partial \eta}{\partial x_i}(z) = 0, \quad 1 \leq i \leq k.$$  

It is trivial to check that $U(z) \geq 1$ and $\frac{\partial \eta}{\partial x_i}(z) \geq 1$ for all $i$ with $1 \leq i \leq n$. Thus, if we define

$$\gamma_i := \frac{1}{n} \frac{\mu'_{\ell,n}(z_i)}{\mu_{\ell,n}(z_i)} \frac{\partial \eta}{\partial x_i}(z), \quad 1 \leq i \leq k,$$

then $\gamma_1 = \cdots = \gamma_k = -\lambda'/f(z)$. Note that

$$\gamma_i = \frac{\sum_{j=0}^{\ell-1} (n-1)_j z_i^j}{\mu_{\ell,n}(z_i) \sum_{j=0}^{\ell-1} \sum_{I \subseteq [k]-i, |I|=j} z_I}, \quad 1 \leq i \leq k.$$  

If $j, k$ are non-negative integers, define

$$Z_{j,k} := \sum_{I \subseteq [k]\{1,2\}, |I|=j} z_I.$$  

Suppose that $z_3, \ldots, z_k$ are fixed. We will show that the equality $\gamma_1 = \gamma_2$ holds if and only if $z_1 = z_2$. By symmetry, this will then prove that $z_1 = \cdots = z_k$, completing the proof that $z$ is a $k$-diagonal point.

Observe that we can write

$$\gamma_1 = \frac{\sum_{j=0}^{\ell-1} (n-1)_j z_1^j}{\mu_{\ell,n}(z_1) \left( (1 + z_2) \sum_{j=0}^{\ell-2} Z_{j,k} + Z_{\ell-1,k} \right)}, \quad (3.3)$$

and

$$\gamma_2 = \frac{\sum_{j=0}^{\ell-1} (n-1)_j z_2^j}{\mu_{\ell,n}(z_2) \left( (1 + z_1) \sum_{j=0}^{\ell-2} Z_{j,k} + Z_{\ell-1,k} \right)}. \quad (3.4)$$

Comparing (3.3) and (3.4), it is now clear that $\gamma_1 = \gamma_2$ if and only if $\Gamma_{\ell,k}(z_1) = \Gamma_{\ell,k}(z_2)$, where

$$\Gamma_{\ell,k}(y) := \frac{\sum_{j=0}^{\ell-1} (n-1)_j y^j \cdot \left( (1 + y) \sum_{j=0}^{\ell-2} Z_{j,k} + Z_{\ell-1,k} \right)}{\mu_{\ell,n}(y)}. \quad (3.5)$$
To conclude, we show that $\Gamma_{\ell,k}(y)$ is strictly decreasing on $[0, \infty)$ and, hence, one-to-one. Note that
\[
\Gamma_{\ell,k}(y) = \Gamma_{\ell,k}(y) \cdot \left( A + \frac{B}{1 + y} \right),
\]
where
\[
\Gamma_{\ell,k}(y) := \frac{\sum_{j=0}^{\ell-1} \binom{n-1}{j} y^j \cdot (1 + y)}{\mu_{\ell,n}(y)}.
\]
and where $A, B > 0$ are constants. We show that $\Gamma_{\ell,k}(y) < 0$. This implies $\Gamma_{\ell,k}(y)$ is strictly decreasing and hence $\Gamma_{\ell,k}(y)$ is strictly decreasing as well. Simple algebra shows that
\[
\Gamma_{\ell,k}(y) = 1 - \frac{(n-1)^\ell}{\sum_{j=0}^{\ell-1} \binom{n}{j} y^j},
\]
and thus
\[
\Gamma'_{\ell,k}(y) = \left( n - 1 \right) \cdot \frac{-\ell y^{\ell-1} \sum_{j=0}^{\ell} \binom{n}{j} y^j + y^{\ell} \sum_{j=0}^{\ell} j \binom{n}{j} y^{j-1}}{\mu_{\ell,n}(y)^2}.
\]
The numerator simplifies to
\[
-z^\ell \sum_{j=0}^{\ell} (\ell - j) \binom{n}{j} y^j < 0,
\]
and the proof is complete.

3.3 Majorization and polynomial inequalities

In this section we use a partial order on polynomials in order to prove Lemma 3.2. Let $f(y) = \sum_{n \geq 0} c(n) y^n$ and $g(y) = \sum_{n \geq 0} d(n) y^n$ be polynomials in $y$ with real coefficients. We say that $f$ is dominated by $g$ in the coefficient partial order, denoted $f \sqsubseteq g$, if and only if $c(n) \leq d(n)$ for all $n \geq 0$. If $f \sqsubseteq g$, and $c(n) < d(n)$ for some $n \geq 0$, then we denote this by $f \subset g$. We write $0 \sqsubseteq f$ if and only if $f$ has all coefficients non-negative and $0 \subset f$ if and only if $f$ has all coefficients non-negative and at least one coefficient positive.

**Proposition 3.4.** If $a > b \geq 0$ are integers, then
\[
\mu_{\ell,a}(y) \mu_{\ell,b}(x) \sqsubseteq \mu_{\ell,a-1}(x) \mu_{\ell,b+1}(x).
\]
We have $\mu_{\ell,a}(x) \mu_{\ell,b}(x) \sqsubset \mu_{\ell,a-1}(x) \mu_{\ell,b+1}(x)$ if and only if, additionally, $a \geq b+2$, $a-1 \geq \ell$, and $\ell \geq 1$.

The proof of Proposition 3.4 requires the two following lemmas.
Lemma 3.5. If $A, B, M, N$ are integers with $A \geq B \geq 0$ and $M \geq N \geq 0$, then
\[
\binom{A}{M} \binom{B}{N} \geq \binom{A}{N} \binom{B}{M}.
\]

For these same ranges of parameters, the inequality is strict if and only if $A > B$, $M > N$, $A \geq M$, and $B \geq N$.

Proof. If $A = B$ or $M = N$ then both sides of the inequality are identically equal. If $M > A$, then $M > B$ and both sides are 0. If $N > B$, then $M > B$ and, again, both sides are 0. We may now assume $A > B$, $M > N$, $A \geq M$, and $B \geq N$. Canceling factorials, the desired inequality becomes
\[
(A - N)_{M-N} > (B - N)_{M-N},
\]
where, for non-negative integers $n$ and real numbers $\alpha$, $(\alpha)_n := \prod_{i=0}^{n-1}(x - i)$ is the falling factorial. Since $(\alpha)_n > (\beta)_n$ if $\alpha > \beta \geq n - 1 \geq 0$, we are done. \qed

Lemma 3.6. Suppose that $f_1, f_2, g_1, g_2$ are polynomials. If $0 \nsubseteq f_1 \nsubseteq f_2$ and $0 \nsubseteq g_1 \nsubseteq g_2$, then $f_1g_1 \nsubseteq f_2g_2$. If, additionally, $f_1 \nsubseteq f_2$ and $0 \nsubseteq g_2$, then $f_1g_1 \nsubseteq f_2g_2$.

Proof. Denoting the coefficients by $f_i = \sum_{j \geq 0} f_{i,j} y^j$ and $g_i = \sum_{j \geq 0} g_{i,j} y^j$ for $i = 1, 2$, the conditions of the lemma state that $0 \leq f_{1,j} \leq f_{2,j}$ and $0 \leq g_{1,j} \leq g_{2,j}$ for all $j \geq 0$. Writing their products as $f_1g_1 = \sum_{j \geq 0} a_j y^j$ and $f_2g_2 = \sum_{j \geq 0} b_j y^j$, the coefficients then satisfy
\[
a_j = \sum_{j,k \geq 0, j+k=l} f_{1,j}g_{1,k} \leq \sum_{j,k \geq 0, j+k=l} f_{2,j}g_{2,k} = b_l,
\]
since $f_{1,j}g_{1,k} \leq f_{2,j}g_{2,k}$ for all $j, k \geq 0$.

If there is additionally some pair $j, k$ such that $0 \leq f_{1,j} < f_{2,j}$ and $0 < g_{2,k}$, then $f_{1,j}g_{1,k} < f_{2,j}g_{2,k}$, and therefore the stronger conclusion $a_{j+k} < b_{j+k}$ holds. \qed

Proof of Proposition 3.4. By definition, proving that $\mu_{\ell,a}(y)\mu_{\ell,b}(y) \nsubseteq \mu_{\ell,a-1}(y)\mu_{\ell,b+1}(y)$ is the same as proving that for each $0 \leq m \leq 2\ell$ we have
\[
\sum_{d=M'}^M \binom{a}{d} \binom{b}{m-d} \leq \sum_{d=M'}^M \binom{a-1}{d} \binom{b+1}{m-d}; \tag{3.6}
\]
where $M' := \max\{0, m - \ell\}$ and $M := \min\{\ell, m\}$. Likewise, the stronger condition that $\mu_{\ell,a}(y)\mu_{\ell,b}(y) \nsubseteq \mu_{\ell,a-1}(y)\mu_{\ell,b+1}(y)$ is equivalent to additionally proving that there is an $m$ with $0 \leq m \leq 2\ell$ for which (3.6) is strict. Applying Pascal’s identity $\binom{a}{d} = \binom{a-1}{d} + \binom{a-1}{d-1}$ to the left-side and $\binom{b+1}{m-d} = \binom{b}{m-d} + \binom{b}{m-d-1}$ to the right, and then cancelling like terms, (3.6) becomes
\[
\sum_{d=M'}^M \binom{a-1}{d} \binom{b}{m-d} \leq \sum_{d=M'}^M \binom{a-1}{d} \binom{b}{m-d-1}; \tag{3.7}
\]
Furthermore, after a summation index shift all of the terms but one in (3.7) cancel, leaving only $d = M'$ on the left-side and $d = M$ on the right:

$$\left( \frac{a - 1}{M' - 1} \right) \left( \frac{b}{m - M'} \right) \leq \left( \frac{a - 1}{M} \right) \left( \frac{b}{m - M - 1} \right).$$

If $M' = 0$, then this inequality is trivially satisfied, and thus so is (3.6). We therefore need only consider the case that $M' = m - \ell$. This means that $m \geq \ell$, so in this case $M = \ell$, and (3.6) is finally equivalent to the inequality

$$\left( \frac{a - 1}{m - \ell - 1} \right) \left( \frac{b}{\ell} \right) \leq \left( \frac{a - 1}{\ell} \right) \left( \frac{b}{m - \ell - 1} \right).$$

We now apply Lemma 3.5 with $A = a - 1$, $B = b$, $M = \ell$, $N = m - \ell - 1$ to complete the proof. The inequality easily follows. It is also easy to see that (3.6) is strict in the cases where $a \geq b + 2$, $a - 1 \geq \ell$, and $2\ell \geq m \geq \ell + 1$ (and hence $\ell \geq 1$).

\[Q.E.D.\]

Remark 3.7. Proposition 3.4 (and its proof) can also be interpreted combinatorially. In particular, consider two rows consisting of $a$ and $b$ square cells, respectively. The $y^m$ coefficient in

$$\sum_{i=0}^{\ell} \binom{a}{i} y^i \cdot \sum_{j=0}^{\ell} \binom{b}{j} y^j$$

is the number of ways of marking exactly $m$ of the cells subject to the restriction that there are at most $\ell$ marked cells in each row, and the result then states that if $a > b$, then there are at least as many ways to mark two rows of length $a - 1$ and $b + 1$ subject to the same restriction.

Corollary 3.8. If $y \geq 0$, $\lambda_1, \lambda_2 \in \Delta(m)$ each have integer coordinates, and $\lambda_1 \succeq \lambda_2$ then

$$\prod_{i=1}^{m} \mu_{\ell, \lambda_1(i)}(y) \supseteq \prod_{i=1}^{m} \mu_{\ell, \lambda_2(i)}(y).$$

Proof. Suppose $\lambda_1 \neq \lambda_2$. By definition, there must then be two indices $1 \leq \alpha < \beta \leq m$ such that $\lambda_1(\alpha) > \lambda_2(\alpha)$ and $\lambda_1(\beta) < \lambda_2(\beta)$. Define $\lambda'_1$ by setting

$$\lambda'_1(\alpha) := \lambda_1(\alpha) - 1, \quad \lambda'_1(\beta) := \lambda_1(\beta) + 1,$$

and $\lambda'_1(i) := \lambda_1(i)$ for all $i \neq \alpha, \beta$. Importantly, it is still true that $\lambda'_1$ majorizes $\lambda_2$.

Noting that $\lambda_1(\alpha) > \lambda_2(\alpha) \geq \lambda_2(\beta) > \lambda_1(\beta)$, Proposition 3.4 now states that

$$\mu_{\ell, \lambda_1(\alpha)}(y) \mu_{\ell, \lambda_1(\beta)}(y) \supseteq \mu_{\ell, \lambda'_1(\alpha)}(y) \mu_{\ell, \lambda'_1(\beta)}(y)$$

which, combined with Lemma 3.6, implies that

$$\prod_{i=1}^{m} \mu_{\ell, \lambda_1(i)}(x) \supseteq \prod_{i=1}^{m} \mu_{\ell, \lambda'_1(i)}(x).$$

(3.8)
If $\lambda_1 = \lambda_2$, then (3.8) gives the statement of the corollary. Otherwise, the above procedure is repeated (a finite number of steps) until this is the case.

Applying this result with the partitions $\lambda_1 = n^k$ and $\lambda_2 = k^n$ will finally complete the proof of Lemma 3.2.

**Proof of Lemma 3.2.** Corollary 3.8 implies $\mu_{\ell,n}(y)^k \subseteq \mu_{\ell,k}(y)^n$. Since this partial order requires that all coefficients be dominated, this immediately implies that $\mu_{\ell,n}(y)^k \leq \mu_{\ell,k}(y)^n$ for all $y \geq 0$. Clearly, if $k = n$, $\ell = 0$, $\ell = n$, or $x = 0$, then $\mu_{\ell,n}(x)^k = \mu_{\ell,k}(x)^n$. It therefore remains to be shown that the inequality is strict if $1 \leq \ell \leq n - 1$, $k \leq n - 1$ and $x > 0$.

Proposition 3.4 implies that $\mu_{\ell,n} \mu_{\ell,0} \subseteq \mu_{\ell,n-1} \mu_{\ell,1}$. Following the proof method of Proposition 3.4, we introduce the dummy term $\mu_{\ell,0} = 1$ and find

$$\mu_{\ell,n}^k = \mu_{\ell,n}^{k-1}(\mu_{\ell,n} \mu_{\ell,0}) \subseteq \mu_{\ell,n}^{k-1}(\mu_{\ell,n-1} \mu_{\ell,1}) \subseteq \mu_{\ell,k}^n.$$

The second relation follows from Lemma 3.6, and the third follows from Corollary 3.8. Since $\mu_{\ell,n}^k \subseteq \mu_{\ell,k}^n$ and $x > 0$, we conclude that $\mu_{\ell,n}(x)^k < \mu_{\ell,k}(x)^n$. 


4 Inequalities for sums of Bernoulli random variables

In this brief section we describe the relationship between our quasi-mean inequalities in Theorem 1.3 and the distributions of sums of Bernoulli random variables.

**Proof of Theorem 1.4.** The inequality is trivial if $\ell = n$, so we henceforth assume that $\ell < n$. Furthermore, if $p_i = 1$ for some $i$ with $1 \leq i \leq n$, then $P(C_i \leq \ell) = 0$ and the inequality is again trivially true. We therefore also assume that $p_i \in [0,1)$ for each $i$.

All of the events $\{C_j \leq \ell\}$ are independent, and their individual probabilities are given by

$$P(C_j \leq \ell) = \sum_{0 \leq m \leq \ell} \left( \begin{array}{c} n \\ m \end{array} \right) p_j^m (1 - p_j)^{n-m}.$$  

Thus

$$P\left( \max_{1 \leq j \leq n} \{C_j\} \leq \ell \right) = \prod_{j=1}^{\ell} \sum_{0 \leq m \leq \ell} \left( \begin{array}{c} n \\ m \end{array} \right) p_j^m (1 - p_j)^{n-m}. \quad (4.1)$$

Similarly, the events $\{R_i \leq \ell\}$ are also independent, and their probabilities are given by

$$P(R_i \leq \ell) = \sum_{0 \leq m \leq \ell} \sum_{I \subseteq [n]} \prod_{j \in I} p_j \prod_{j \notin I} (1 - p_j),$$

so

$$P\left( \max_{1 \leq i \leq n} \{R_i\} \leq \ell \right) = \left( \sum_{0 \leq m \leq \ell} \sum_{I \subseteq [n]} \prod_{j \in I} p_j \prod_{j \notin I} (1 - p_j) \right)^n. \quad (4.2)$$
Dividing (4.1) and (4.2) by $\prod_{j=1}^{n}(1 - p_j)$, we see that the desired inequality is equivalent to the left inequality from Theorem 1.3 with $x_i = p_i/(1 - p_i)$ for $1 \leq i \leq n$. □

References


