Analytical Lower Bounds for the Probability of at Least m-Out-Of-n Events Occurring

Songlan Nie, Jun Yang

University of Oxford

A thesis submitted for the degree of

MMath in Mathematics and Statistics

Trinity 2023

Abstract

In the paper, we discussed analytical lower bounds for the probability of at least m-out-of-n events occurring, $\mathbb{P}(\mu \geqslant m)$, where μ is the number of events occurring, given information about individual probabilities and aggregated/disaggregated joint probabilities up to t events. In particular, we generalised the work by Prékopa and Gao on $\mathbb{P}(\mu \geqslant 1)$ and derived new analytical lower bounds for the two cases $m = 2, t = 2$ and $m = 3, t = 3$. The derivations were carried out through either probabilistic inequalities or linear programmings (LPs). We considered the particular relaxation of LP used in the derivation of the Yang-Alajaji-Takahara (YAT) bound and applied the additional constraints on the joint probability of n events to strengthen the bounds derived. For the $m = 2, t = 2$ case, we derived lower bounds using the Cauchy-Schwarz inequality and LPs with two classes of disaggregated partial information and managed to show that the two bounds are in fact identical although the latter is usually sharper as shown in the literature. For the $m = 3, t = 3$ case, we constructed a bound using LPs with three classes of disaggregated partial information. Since our bounds are derived using LPs with disaggregated information and additional constraints, they are sharper than the corresponding bounds with aggregated partial information by Boros and Prékopa by construction.

Contents

Chapter 1

Introduction

1.1 Motivation

It is a common practice to model the reliability of a stochastic system using the probability of at least m-out-of-n events occurring [1]. While there are some closed-form formulas for it, they are usually computationally prohibitive to use in practice. Hence, it is important to derive bounds for these probabilities. For example, there are n machines running simultaneously and each machine has some probability of breaking down. Assume that we need at least m machines running to keep the system working. Then the lower bound for the probability of having at least m-out-of-n working machines provides a conservative estimate of the probability that guarantees the running of the system.

Let $A_1, \ldots, A_n$ be a sequence of n events on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, where n is a fixed positive integer and μ is the number of events occurring. Here, μ is an integer-valued random variable that takes values between 0, where no event occurs, and n, where all n events occur. There are some formulas that compute $\mathrm{P}(\mu \geqslant m)$ exactly. For the case $m = 1$, we have the *inclusion-exclusion formula* [2]

$$\mathbb{P}(\mu \geqslant 1) = \mathbb{P}\left(A_1 \cup \cdots \cup A_n\right) = S_1 - S_2 + \cdots + (-1)^{n-1} S_n \tag{1.1}$$

where

$$S_t = \sum_{1 \leqslant i_1 < \cdots < i_t \leqslant n} \mathbb{P}\left(A_{i_1} \cap \cdots \cap A_{i_t}\right). \tag{1.2}$$

There are also formulas for a general m using the *Jordan's formula* [3]:

$$\mathrm{P}(\mu \geqslant m) = \sum_{t=m}^{n} (-1)^{t-m} \binom{t-1}{m-1} S_t. \tag{1.3}$$

However, when n is large, it is impractical to use these formulas [1] since the number of terms involved grows exponentially in n. In practice, we are usually given $S_1, \ldots, S_t$ for $t \ll n$. These insufficient constraints on S_t's are called *partial information*. In this circumstance, the best we can achieve is to determine bounds for $\mathrm{P}(\mu \geqslant m)$.

In the literature, there have been attempts in deriving both upper and lower bounds for $\mathrm{P}(\mu \geqslant m)$ [4, 2, 5]. In this paper, we will focus on finding a lower bound ℓ such that $\mathbb{P}(\mu \geqslant m) \geqslant \ell$ using a few S_t's. Note that these S_t's do not give us information about a particular event A_i but rather in terms of a sum of individual probabilities. We call these S_t's the *aggregated constraints* and the corresponding bound finding problems the *aggregated problems*.

If instead we are given information about individual probabilities specific to A_i, for example $\{\mathbb{P}(A_i)\}_{i=1}^n$ and $\{\sum_{1\leqslant j\leqslant n: j\neq i} \mathbb{P}(A_i \cap A_j)\}_{i=1}^n$, these constraints are called *disaggregated constraints* and the corresponding problems are called *disaggregated problems*. We can also formally define the corresponding disaggregated partial information as follows [6]:

$$S_t^i := \sum_{\substack{1\leqslant i_1 < \cdots < i_t \leqslant n \\ \{i_1, \ldots, i_t\} \ni i}} \mathbb{P}\left(A_{i_1} \cap \cdots \cap A_{i_t}\right). \tag{1.4}$$

This is the sum of probabilities of of intersections of t events, one of which is A_i.

There are in general two types of lower bounds in the literature [5]: bounds derived from probabilistic inequalities and bounds derived using linear programmings (LPs). Classical examples for the first type include bounds derived from the Cauchy-Schwarz inequality and the Bonferroni inequality [5]. For simplicity, we will call a bound derived from the Cauchy-Schwarz (CS) inequality a CS type bound and the one derived from the Bonferroni inequality a Bonferroni type bound. Examples for the former type include one bound for the aggregated case mentioned in [5] and the so-called D. de Caen (DC) bound [7] for the corresponding disaggregated case. One result for the Bonferroni type bound is given as follows [8]

$$P\left(\bigcup_{i=1}^n A_i\right) \geqslant \sum_{i=1}^n P\left(A_i\right) - \sum_{1\leqslant i<j\leqslant n} P\left(A_i \cap A_j\right). \tag{1.5}$$

While lower bounds through probabilistic inequalities might be easier to derive, they are usually loose (small). Bounds derived using LPs are generally sharper and sometimes even optimal. Formally, an *LP problem* is an optimisation problem of a linear function subject to linear equality and inequality constraints. By *sharper lower* *bounds*, we mean the bound is larger and thus provides a better approximation for $\mathrm{P}(\mu \geqslant m)$. By *optimal lower bounds*, we mean there is no other sharper bound possible.

In general, LPs cannot be solved explicitly and we have to resort to numerical software, making the bound more cumbersome to use. Fortunately, LPs that can be solved explicitly give us the benefits of both types of bounds: sharp and easy to implement. There are a number of existing works that give analytical/closed-form lower bounds. Works by Dawson and Sankoff [9] gave the optimal bound for $\mathrm{P}(\mu \geqslant 1)$, named the DS bound, using the aggregated partial information S_1 and S_2. Boros and Prékopa [10] provided a comprehensive derivation of the bounds for $\mathrm{P}(\mu \geqslant m)$ using $S_1, \dots, S_t$, where closed-form bounds are derived for $t = 2, 3, 4$. In addition, the Kuai-Alajaji-Takahara (KAT) bound [11] is a closed-form bound for $\mathrm{P}(\mu \geqslant 1)$ using the disaggregated partial information $\{\mathbb{P}(A_i)\}_{i=1}^n$ and $\{\sum_{j:j\neq i} \mathbb{P}(A_i \cap A_j)\}_{i=1}^n$, or equivalently using S_1^i and S_2^i for $i = 1, \dots, n$. Yang, Alajaji and Takahara [12] derived a bound that improves the KAT bound further, which we call the YAT bound, by adding constraints on $\mathbb{P}(A_1 \cup \cdots \cup A_n)$ to the LP problem for the KAT bound. Prékopa and Gao [2] applied the results from Boros and Prékopa [10] to the corresponding disaggregated problems for $\mathrm{P}(\mu \geqslant 1)$ using the disaggregated partial information $S_1^i, \dots, S_t^i$ for $i = 1, \dots, n$ with analytical bounds derived for $t = 2, 3, 4$. Boros and Lee [4] gave a systematic review of the construction of primal-dual pair of LPs that induce the bounds for $\mathrm{P}(\mu \geqslant m)$ using both aggregated and disaggregated information. However, there are very limited works on analytical bounds for $\mathrm{P}(\mu \geqslant m)$ using disaggregated partial information with $m \geqslant 2$. Works like [13, 14] discussed analytical bounds for $\mathrm{P}(\mu \geqslant m)$ and $\mathrm{P}(\mu = m)$ in the context of unimodal distributions, i.e. the probability distribution of μ has a single mode. This assumption is somewhat restrictive and the results do not apply to the general situation. We aim to construct bounds that work in a general context in this paper.

In this paper, we made the following contributions. Inspired by the idea of imposing extra constraints from the YAT bound [12], we added the constraints to the bound finding problems for $\mathrm{P}(\mu \geqslant m)$ with $m = 2, 3$. For the $m = 2$ case, we derived lower bounds using the Cauchy-Schwarz inequality and LPs with disaggregated partial information S_1^i, S_2^i for $i = 1, \dots, n$ and we managed to show that the two bounds are in fact identical. For the $m = 3$ case, we constructed a bound using LPs with disaggregated partial information S_1^i, S_2^i, S_3^i for $i = 1, \dots, n$. Since our bounds are derived using LPs with disaggregated information (thus more information) and additional constraints, they are sharper than bounds [10, 2] by construction. The bounds derived can provide foundations for future works in analytical bounds for $\mathrm{P}(\mu \geqslant m)$ with a general m.

1.2 Problem framework

Before introducing the framework of deriving lower bounds of probabilities, consider a simple setup of three events A_1, A_2, A_3 and the following Venn diagram.

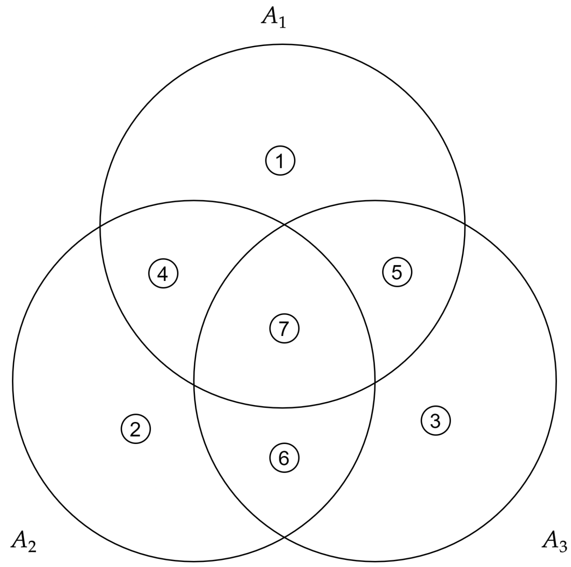

Figure 1.1: Venn Diagram

Here, events A_1, A_2, A_3 are represented by the three circles. The overlaps between different circles represent multiple events happening. In this way, the space is divided into seven small indivisible regions, each associated with a probability for the corresponding event. If we define $p_i := \mathbb{P}(\textcircled{i})$ for $i = 1, \ldots, 7$. Then, we have the following basic result:

$$\begin{aligned}
\mathbb{P}(A_1) &= p_1 + p_4 + p_5 + p_7, \\
\mathbb{P}(A_2) &= p_2 + p_4 + p_6 + p_7, \\
\mathbb{P}(A_3) &= p_3 + p_5 + p_6 + p_7.
\end{aligned} \tag{1.6}$$

We can also consider the intersection of events. For example,

$$\mathbb{P}(A_1 \cap A_2) = p_4 + p_7. \tag{1.7}$$

Now we can turn to our problem of interest. If at least one out of three events occurs, then any one out of the seven regions will satisfy the condition. Hence,

$$\mathrm{P}(\mu \geqslant 1) = \mathbb{P}(A_1 \cup A_2 \cup A_3) = \sum_{i=1}^{7} p_i. \tag{1.8}$$

Similarly, for the two-out-of-three and the three-out-of-three cases, we have

$$\mathrm{P}(\mu \geqslant 2) = \mathbb{P}((A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3)) = \sum_{i=4}^{7} p_i, \tag{1.9}$$

$$\mathrm{P}(\mu \geqslant 3) = \mathbb{P}(A_1 \cap A_2 \cap A_3) = p_7. \tag{1.10}$$

Crucially, we see that $\mathrm{P}(\mu \geqslant m)$ can be decomposed into a sum of probabilities of small non-intersecting regions so that it becomes much more tractable to work with.

Now we will formally introduce the framework, defining partial information and lower bounds as follows.

Definition 1.2.1. *(Partial information) Let $A_1, \ldots, A_n$ be a sequence of n events. Let H be the space of functions of $\mathbb{P}(A_{i_1}), \mathbb{P}(A_{i_1} \cap A_{i_2}), \ldots, \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_n})$ denoted by $\eta_s : [0,1]^{n+\binom{n}{2}+\ldots+\binom{n}{n}} \to \mathbb{R}^s$ where $s \in \mathbb{Z}^{>0}$ and the range of all such functions as Θ. Then the partial information of the n events is defined as the function values $\theta \in \Theta$ taken by the function $\eta_s \in H$ [5].*

Remark 1.2.2. *This definition is a generalisation of the formation in the reference [5], in which the space of functions only includes functions of $\mathbb{P}(A_{i_1}), \mathbb{P}(A_{i_1} \cap A_{i_2})$.*

Example 1.2.3. *Here are two examples of partial information [5] where $s = n$ and $s = 2$, respectively.*

$$\theta = \left(\mathbb{P}\left(A_1\right), \mathbb{P}\left(A_2\right), \ldots, \mathbb{P}\left(A_n\right)\right), \tag{1.11}$$

$$\theta = \left(\sum_{i=1}^{n} \mathbb{P}\left(A_i\right), \sum_{i,j=1}^{n} \mathbb{P}\left(A_i \cap A_j\right)\right). \tag{1.12}$$

The first one gives information about the probabilities of events A_i's for each $i = 1, \ldots, n$ while the second one specifies the sum of the probability of each events and the intersection of any two events.

Note that for large n, both θ's give us some information about $\mathbb{P}(\cup_{i=1}^{n} A_i)$ but are not sufficient to determine the union probability exactly, justifying the name "partial" information.

Remark 1.2.4. *Note that the definition of partial information does not exclude the case where we can determine $\mathbb{P}(\mu \geqslant m)$ exactly when s is sufficiently large. However, the purpose of the research is to determine a lower bound for $\mathbb{P}(\mu \geqslant m)$ given limited partial information. Hence, we may assume $s \ll n$ so it is not possible to determine $\mathbb{P}(\mu \geqslant m)$ exactly.*

Definition 1.2.5. *(Lower bounds for the probability of events) A lower bound for the probability of at least m-out-of-n events occurring is a function of θ, $\ell(\theta)$ such that*

$$\mathbb{P}(\mu \geqslant m) \geqslant \ell(\theta) \tag{1.13}$$

for any set of events $\{A_i\}_{i=1}^n$ that the value of η_s for given $\{A_i\}_{i=1}^n$ equals to θ. Let $\mathscr{L}_\Theta^m$ denote the set of all lower bounds on $\mathbb{P}(\mu \geqslant m)$ that are functions of only θ [5].

We want this lower bound to be as large as possible to be useful. The formal way to call the largest of such bound is the optimal lower bound.

Definition 1.2.6. *(Optimality) A lower bound $\ell^\star \in \mathscr{L}_\Theta$ is optimal in $\mathscr{L}_\Theta$ if $\ell^\star(\theta) \geqslant \ell(\theta)$ for all $\theta \in \Theta$ and $\ell \in \mathscr{L}_\Theta$ [5].*

Another seemingly unrelated concept is the achievability of the lower bound. By achievability, we mean it is possible to construct a sequence of events $\{A_i\}_{i=1}^n$ that attain the lower bound. In fact, these two concepts are equivalent.

Definition 1.2.7. *(Achievability) A lower bound $\ell \in \mathscr{L}_\Theta$ is achievable if for every $\theta \in \Theta$,*

$$\inf_{A_1,\dots,A_n} \mathbb{P}(\mu \geqslant m) = \ell(\theta). \tag{1.14}$$

where the infimum ranges over all sets of events $\{A_i\}_{i=1}^n$ constrained by θ [5].

Lemma 1.2.8. *(Equivalence of optimality and achievability) A lower bound $\ell^\star \in \mathscr{L}_\Theta$ is optimal in $\mathscr{L}_\Theta$ if and only if it is achievable.*

The proof for this lemma is not central to our discussion and it can be found in the reference [5].

Note that the event $\{\mu \geqslant m\}$ is very difficult to work with directly. Hence, we will introduce some concepts to decompose this event into unions of smaller event.

Definition 1.2.9. *(Elementary/simple event) An elementary/simple event is an event that cannot be written as the union of two events with positive probabilities [15].*

Remark 1.2.10. *On the contrary, a compound event is the union of different elementary events [15].*

Example 1.2.11. *Consider rolling an unbiased die. We can define the event space as the power set of* $\{1,2,3,4,5,6\}$. *Then* $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$ *are elementary events since they cannot be written as a union of two events with positive probability. On the contrary,* $\{1,2\}$ *is a compound event as* $\{1,2\} = \{1\} \cup \{2\}$ *and* $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \frac{1}{6} \neq 0$.

Example 1.2.12. *We return to our earlier example involving three events* A_1, A_2, A_3. *All seven events* 1, …, 7 *are elementary while* $A_1 \cap A_2 =$ 4 $\cup$ 7 *is compound.*

Remark 1.2.13. *One important property of elementary events is that different elementary events are disjoint and their probabilities are additive. Take two elementary events* $\omega_1, \omega_2 \in \mathscr{F}$ *with* $\omega_1 \neq \omega_2$. *Then* $\omega_1 \cap \omega_2 = \varnothing$ *since otherwise* $\omega_1 = (\omega_1 \cap \omega_2^c) \cup (\omega_1 \cap \omega_2)$ *and* ω_1 *is not elementary. Hence, we can decompose the probability of a compound event into a sum of probabilities of elementary events, making the bound finding problems tangible.*

In order to classify different elementary events, we introduce the concept of the degree of an elementary event in the following.

Definition 1.2.14. *(Degree of an elementary event) For each elementary event* $\omega \in \mathscr{F}$, *the degree of* ω *is defined as the number of event* A_i*'s that contain* ω, *i.e.* $\deg(\omega) := \#\{A_i : \omega \subseteq A_i\}$.

Remark 1.2.15. *For a compound event, its degree is not well defined since it can be written as a union of different elementary events that are contained in different event* A_i*'s. Hence, there exists some event* A_i *that only contains some elementary events in that compound event but not all of them, which contradicts with our definition, since we need the event to be a subset of* A_i. *To avoid any confusion, when we specify the degree of an event, we assume that this event is elementary.*

Remark 1.2.16. *From the definition, if* k *events occur, then any degree* k *event occurs since a degree* k *event is contained in exactly* k *events. Likewise, if at least* m *events occur, then any event of degree at least* m *occurs.*

Definition 1.2.17. *(Probability of degree* k *events) The probability of degree* k *events is defined as* $a(k) := \mathbb{P}\left(\{\omega \subseteq \bigcup_{i=1}^{n} A_i : \deg(\omega) = k\}\right)$. *The probability of degree* k *events contained in* A_i *is defined as* $a_i(k) := \mathbb{P}\left(\{\omega \subseteq A_i : \deg(\omega) = k\}\right)$.

Remark 1.2.18. *Note that a degree k event is contained in k different event A_i's. Hence, summing the probability of degree k events contained in A_i over i will count each degree k event exactly k times. Hence, $ka(k) = \sum_{i=1}^{n} a_i(k) \iff a(k) = \sum_{i=1}^{n} \frac{a_i(k)}{k}$.*

We can specify the values of $a(k)$'s and $a_i(k)$'s as partial information. This notation is different from joint probabilities. In the literature, partial information is usually given as joint probabilities up to t events.

One commonly used notations for disaggregated partial information in the literature is S_t^i as defined in the previous section. Now we can specify disaggregated partial information in terms of either $a_i(k)$ or S_t^i. There is also a relationship between S_t^i and $a_i(k)$ which is proved in [6]:

$$\sum_{k=t}^{n} \binom{k-1}{t-1} a_i(k) = S_t^i. \tag{1.15}$$

We will prove that the two formulations of partial information given in terms of $a_i(k)$ and S_t^i are equivalent.

Lemma 1.2.19. *(Equivalence of partial information) The formulations of partial information in terms of $S_1^i, \dots, S_t^i$ and $\sum_{k=1}^{n} a_i(k), \dots, \sum_{k=1}^{n} k^{t-1} a_i(k)$ for $i = 1, \dots, n$ are equivalent.*

Proof. Note first that $\binom{k-1}{t-1}$ is a polynomial in k of degree $t-1$ with coefficient $\frac{1}{t-1}$ for the highest order term k^{t-1}. Then, in matrix notation $Pa = s$, we have

$$\begin{bmatrix} 1 & & & & \\ * & \frac{1}{2} & & & \\ \vdots & * & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ * & \cdots & \cdots & * & \frac{1}{t-1} \end{bmatrix} \begin{bmatrix} \sum_{k=1}^{n} a_i(k) \\ \sum_{k=1}^{n} k a_i(k) \\ \vdots \\ \vdots \\ \sum_{k=1}^{n} k^{t-1} a_i(k) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_i(k) \\ \sum_{k=1}^{n} (k-1) a_i(k) \\ \vdots \\ \vdots \\ \sum_{k=1}^{n} \binom{k-1}{t-1} a_i(k). \end{bmatrix} \tag{1.16}$$

where $*$ represents some matrix entries we are not interested in.

Note that the RHS is equal to $\begin{bmatrix} S_1^i, & S_2^i, & \dots, & S_t^i \end{bmatrix}^\top$. Then we have a linear system that expresses $S_1^i, \dots, S_t^i$ in terms of $\sum_{k=1}^{n} a_i(k), \dots, \sum_{k=1}^{n} k^{t-1} a_i(k)$. Note that the matrix P is a lower triangular matrix with diagonal entries $1, \frac{1}{2}, \dots, \frac{1}{t-1}$. Then $\det P = \frac{1}{(t-1)!} \neq 0$ and thus P is invertible. Hence, the linear system has a unique solution. Similarly, we can also express $\sum_{k=1}^{n} a_i(k), \dots, \sum_{k=1}^{n} k^{t-1} a_i(k)$ in terms of $S_1^i, \dots, S_t^i$ using P^{-1}. Hence, there is a unique correspondence between $S_1^i, \dots, S_t^i$ and $\sum_{k=1}^{n} a_i(k), \dots, \sum_{k=1}^{n} k^{t-1} a_i(k)$, i.e. the two formulations are equivalent. $\square$

Remark 1.2.20. *Using this lemma, we can work with the formulation of partial information* $\sum_{k=1}^{n} a_i(k), \ldots, \sum_{k=1}^{n} k^{t-1} a_i(k)$ *and it can still be compatible with the literature.*

With the new notation, we can express $\mathbb{P}(\mu \geqslant m)$ in terms of $a_i(k)$.

Proposition 1.2.21. *(Probability of at least m-out-of-n events occurring) The probability of at least m-out-of-n events occurring is given by*

$$\mathbb{P}(\mu \geqslant m) = \sum_{k=m}^{n} a(k) = \sum_{i=1}^{n} \sum_{k=m}^{n} \frac{a_i(k)}{k}. \tag{1.17}$$

Proof.

$$\begin{aligned}
\mathbb{P}(\mu \geqslant m) &= \mathbb{P}\left(\left\{\omega \subseteq \bigcup_{i=1}^{n} A_i : m \leqslant \deg(\omega) \leqslant n\right\}\right) \\
&= \mathbb{P}\left(\bigcup_{k=m}^{n} \left\{\omega \subseteq \bigcup_{i=1}^{n} A_i : \deg(\omega) = k\right\}\right) \\
&= \sum_{k=m}^{n} \mathbb{P}\left(\left\{\omega \subseteq \bigcup_{i=1}^{n} A_i : \deg(\omega) = k\right\}\right) \\
&= \sum_{k=m}^{n} a(k) = \sum_{k=m}^{n} \sum_{i=1}^{n} \frac{a_i(k)}{k} = \sum_{i=1}^{n} \sum_{k=m}^{n} \frac{a_i(k)}{k}.
\end{aligned} \tag{1.18}$$

The first two equalities follow from the definitions of μ and Remark 1.2.16. The third equality follows from the fact that elementary events are disjoint as in Remark 1.2.13. The last three equalities follow from Remark 1.2.18. □

Remark 1.2.22. *This proposition expresses the probability of concern in terms of the sum of the probability of events of particular degrees* $a(k)$ *or* $a_i(k)$*. It is much easier to find bounds for the latter so it suffices to work with the sum directly.*

In the next few sections, we will introduce a number of lower bounds constructed through linear programming (LP) problems. In order to solve those LP problems, we need some important definitions and lemmas to characterise their solutions.

Definition 1.2.23. *(Active constraints) Constraints* $g(x) \geqslant 0$ *and* $h(x) = 0$ *are active at* x^* *if* $g(x^*) = 0$ *and* $h(x^*) = 0$ *[16].*

Example 1.2.24. *Consider an optimisation problem with a decision variable* $x = (x_1, x_2) \in \mathbb{R}^2$*. Then the constraint* $x_1 + x_2 \geqslant 0$ *is active at* $x^* = (x_1^*, x_2^*)$ *if* $x_1^* + x_2^* = 0$.

Definition 1.2.25. *(Feasible points) x^* for an optimisation problem is a feasible point if it satisfies all the constraints.*

Definition 1.2.26. *(Basic feasible solution) A basic feasible solution $x^* \in \mathbb{R}^n$ is a feasible point such that n out of all active constraints are linearly independent [16].*

Example 1.2.27. *Consider an optimisation problem with a decision variable $x = (x_1, x_2) \in \mathbb{R}^2$ and constraints $x_1, x_2 \geqslant 0$ and $x_1 + x_2 = 2$. Since $n = 2$, we need to have two linearly independent active constraints to deduce a basic feasible solution. $x_1 + x_2 = 2$ is already active so either $x_1 \geqslant 0$ or $x_2 \geqslant 0$ is active. If $x_1 \geqslant 0$ is active, then $x_1 = 0$ and $x_1 + x_2 = 2$. These two constraints are linearly independent and we get $(x_1, x_2) = (0, 2)$. Here $x_2 = 2 \geqslant 0$ so this point is feasible. We also have two active linearly independent constraints so $(0, 2)$ is a basic feasible solution. Likewise, if $x_2 \geqslant 0$ is active we obtain the basic feasible solution $(x_1, x_2) = (2, 0)$.*

There might exist multiple feasible solutions to an LP problem that give rise to the same value for the objective function. We can determine any one of them which is enough for our purpose. One important result is that it suffices to find a basic feasible solution to solve an LP.

Lemma 1.2.28. *(Optimal basic feasible solution) For an LP problem, if there exists an optimal solution, then there exists an optimal basic feasible solution [16].*

Using the previous lemma and the definition of basic feasible solutions, we can derive an algorithm for solving an LP. Given some linear equality and inequality constraints, we can first set some inequality constraints to be active so that there are n linearly independent active constraints. Now we have n independent linear equations for a decision variable $x \in \mathbb{R}^n$. Then we can solve for x uniquely and obtain the value of the objective function. Repeating this process and choosing different sets of active constraints, we can compare different values of the objective function and choose the smallest among all.

We will consider a specific case of LP and formalise the above argument into an algorithm.

Algorithm 1.2.29. *Suppose we have an LP for $x \in \mathbb{R}^n$ with m equality constraints with $m < n$ and $x_i \geqslant 0$ for $i = 1, \ldots, n$. From this lemma, we can derive the following algorithm for solving an LP problem:*
1. Set $(n - m)$ x_i's to be zero and solve for x,
2. Compare different x's and find the optimal x^* that minimises the objective function.

Chapter 2

Review of Existing Bounds

2.1 Analytical lower bounds for $\mathbb{P}(\mu \geqslant 1)$

The probability of at least 1-out-of-n events occurring corresponds to the case $m = 1$ in our framework. This is also commonly called the union probability and it receives the most attention in the literature. In the work by Yang et al. [5], a range of analytical bounds are investigated. These bounds mainly fall into two categories. They are either constructed using probabilistic inequalities or through LP problems. Probabilistic inequalities are straightforward to use but the bounds derived from them are relatively loose. On the other hand, the bounds derived from LP problems are at least as sharp as those and sometimes optimal given suitable constraints. In the following few sections, we will discuss these two bounds in details and the reason favouring the bounds derived from LP problems.

2.1.1 Bounds through probabilistic inequalities

Lemma 2.1.1. *(Cauchy-Schwarz inequality) Let* $\mathbf{x}, \mathbf{y}$ *be two elements from an inner product space associated with the inner product* $\langle \cdot, \cdot \rangle$. *Then* $|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leqslant \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle$.

Now we will construct our simplest possible bound derived from the Cauchy-Schwarz inequality. For simplicity, we will call a bound derived from the Cauchy-Schwarz (CS) inequality a CS type bound.

Proposition 2.1.2. *(CS type bound, aggregated constraints) Consider the problem of finding a lower bound* $\ell(\theta)$ *for* $\mathbb{P}(\mu \geqslant 1)$ *with aggregated partial information* $\theta = (\alpha, \beta)$

given in the following:

$$\begin{aligned}
\ell(\theta) \leqslant \mathbb{P}(\mu \geqslant 1) = \sum_{k=1}^{n} a(k) &= \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{a_i(k)}{k} \\
s.t. \sum_{k=1}^{n} ka(k) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_i(k) &= \alpha, \qquad (2.1)\\
\sum_{k=1}^{n} k^2 a(k) = \sum_{i=1}^{n}\sum_{k=1}^{n} ka_i(k) &= \beta.
\end{aligned}$$

Then we can derive the bound $\ell_{CS} := \frac{\alpha^2}{\beta}$, *which we call the CS bound, using the Cauchy-Schwarz inequality.*

Proof. Note that the Euclidean dot product is an inner product. Then by the Cauchy-Schwarz inequality, we have

$$\left(\sum_{k=1}^{n} ka(k)\right)^2 = \left(\sum_{k=1}^{n} \sqrt{a(k)}\sqrt{k^2a(k)}\right)^2 \tag{2.2}$$

$$\leqslant \left(\sum_{k=1}^{n} \left(\sqrt{a(k)}\right)^2\right)\left(\sum_{k=1}^{n} \left(\sqrt{k^2a(k)}\right)^2\right) \tag{2.3}$$

$$= \left(\sum_{k=1}^{n} a(k)\right)\left(\sum_{k=1}^{n} k^2a(k)\right). \tag{2.4}$$

Rearranging, we get

$$\sum_{k=1}^{n} a(k) \geqslant \frac{\left(\sum_{k=1}^{n} ka(k)\right)^2}{\sum_{k=1}^{n} k^2a(k)} = \frac{\alpha^2}{\beta} =: \ell_{CS}. \tag{2.5}$$

□

One improvement made to the CS bound is replacing the aggregated constraints

$$\sum_{i=1}^{n}\sum_{k=1}^{n} a_i(k) = \alpha,\ \sum_{i=1}^{n}\sum_{k=1}^{n} ka_i(k) = \beta \tag{2.6}$$

with the disaggregated constraints

$$\sum_{k=1}^{n} a_i(k) = \alpha_i,\ \sum_{k=1}^{n} ka_i(k) = \beta_i,\ i = 1, \ldots, n. \tag{2.7}$$

Then we have the D. de Caen (DC) bound [7, 5].

Proposition 2.1.3. *(CS type bound, disaggregated constraints) Consider the problem of finding a lower bound $\ell(\theta)$ for $\mathbb{P}(\mu \geqslant 1)$ with disaggregated partial information*

$$\theta = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) \tag{2.8}$$

given in the following:

$$\begin{aligned} \ell(\theta) \leqslant \mathbb{P}(\mu \geqslant 1) = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{a_i(k)}{k} \\ s.t. \sum_{k=1}^{n} a_i(k) = \alpha_i, \; i = 1, \ldots, n, \\ \sum_{k=1}^{n} k a_i(k) = \beta_i, \; i = 1, \ldots, n. \end{aligned} \tag{2.9}$$

Then we can derive the DC bound $\ell_{DC} := \sum_{i=1}^{n} \frac{\alpha_i^2}{\beta_i}$ using the Cauchy-Schwarz inequality.

Proof. Using a similar argument, by the Cauchy-Schwarz inequality, we get

$$\left(\sum_{k=1}^{n} a_i(k) \right)^2 = \left(\sum_{k=1}^{n} \sqrt{\frac{a_i(k)}{k}} \sqrt{k a_i(k)} \right)^2 \tag{2.10}$$

$$\leqslant \left(\sum_{k=1}^{n} \frac{a_i(k)}{k} \right) \left(\sum_{k=1}^{n} k a_i(k) \right). \tag{2.11}$$

Rearranging, we have

$$\sum_{k=1}^{n} \frac{a_i(k)}{k} \geqslant \frac{\left(\sum_{k=1}^{n} a_i(k) \right)^2}{\sum_{k=1}^{n} k a_i(k)} = \frac{\alpha_i^2}{\beta_i}. \tag{2.12}$$

Summing over i gives

$$\sum_{i=1}^{n} \sum_{k=1}^{n} \frac{a_i(k)}{k} \geqslant \sum_{i=1}^{n} \frac{\alpha_i^2}{\beta_i} =: \ell_{DC}. \tag{2.13}$$

□

Remark 2.1.4. *Note that given the disaggregated constraints*

$$\sum_{k=1}^{n} a_i(k) = \alpha_i, \; \sum_{k=1}^{n} k a_i(k) = \beta_i, \; i = 1, \ldots, n, \tag{2.14}$$

we are automatically given the aggregated constraints

$$\sum_{i=1}^{n}\sum_{k=1}^{n} a_i(k) = \sum_{k=1}^{n} \alpha_i =: \alpha, \sum_{i=1}^{n}\sum_{k=1}^{n} a_i(k) = \sum_{k=1}^{n} \beta_i =: \beta. \tag{2.15}$$

Then, applying the previous proposition, we get the bound

$$\ell_{CS} = \frac{\alpha^2}{\beta} = \frac{(\sum_{k=1}^{n} \alpha_i)^2}{\sum_{k=1}^{n} \beta_i}. \tag{2.16}$$

Hence, we may always derive a bound involving aggregate constraints on $a(k)$ using individual constraints on $a_i(k)$. We can also show that the DC bound is sharper than the CS bound. By the Cauchy-Schwarz inequality,

$$\left(\sum_{k=1}^{n} \alpha_i\right)^2 = \left(\sum_{k=1}^{n} \frac{\alpha_i}{\sqrt{\beta_i}}\sqrt{\beta_i}\right)^2 \leqslant \left(\sum_{k=1}^{n} \frac{\alpha_i^2}{\beta_i}\right)\left(\sum_{k=1}^{n} \beta_i\right). \tag{2.17}$$

Rearranging, we get

$$\ell_{CS} = \frac{(\sum_{k=1}^{n} \alpha_i)^2}{\sum_{k=1}^{n} \beta_i} \leqslant \sum_{i=1}^{n} \frac{\alpha_i^2}{\beta_i} = \ell_{DC}. \tag{2.18}$$

This matches our intuition that the DC bound, which uses more information, is better than the CS bound.

2.1.2 Bounds through the construction of LP problems

There is extensive literature on the construction of lower bounds for the probability of at least 1-out-of-n events occurring, commonly known as the union probability, through LP problems. This method produces the best possible bound by construction.

The reason is as follows. Consider the set of possible values S that the the probability of at least 1-out-of-n events occurring can take.

$$S := \left\{ \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{a_i(k)}{k} \text{ subject to linear constraints} \right\}. \tag{2.19}$$

We are concerned about the problem

$$\text{find } \ell(\theta) \leqslant s \quad \forall s \in S. \tag{2.20}$$

Note that S is a set of probabilities so it is bounded below by 0. Hence, the $\inf S$ exists by the Completeness Axiom. Then the largest possible ℓ that satisfies the constraints is $\inf S$ since the infimum of a set is the largest possible lower bound. If the constraints give rise to a closed feasible region in $\mathbb{R}$, then by continuity of the objective function, S is closed in $\mathbb{R}$ so $\min S \in S$. This is true since the constraints are either linear equalities or linear weak inequalities. This shows that LP bounds are the best possible bounds that satisfy the constraints.

However, there is one caveat here. The bounds constructed on the objective function might still not be optimal for $\mathbb{P}(\mu \geqslant 1)$ given the constraints. The reason is that the LP problem is constructed in terms of $a_i(k)$'s but not the actual events $A_1, \ldots, A_n$. These $a_i(k)$'s only satisfy the explicit constraints imposed in the LP problem. There are certain additional implicit constraints that need to be satisfied to be able to construct events $A_1, \ldots, A_n$ that are feasible. One such example is the constraint $\sum_{i=1}^n a_i(k) \geqslant k a_j(k)$ for $j, k = 1, \ldots, n$ that can be imposed on the LP used in the KAT bound [12]. The details will be discussed in the following.

2.1.2.1 DS bound

One of the simplest possible LP bound for $\mathbb{P}(\mu \geqslant 1)$ is the Dawson-Sankoff (DS) bound [9] which involves only two disaggregated constraints. It has the same setup as Proposition 2.1.2 but is sharper than the CS type bound as discussed in [5]. The bound is introduced in the following proposition.

Proposition 2.1.5. *(DS bound) The DS bound is the solution to the following LP problem*

$$
\begin{aligned}
\ell_{DS} := \min_{a(k):k=1,\ldots,n} & \sum_{k=1}^n a(k), \\
s.t. \quad & \sum_{k=1}^n k a(k) = \sum_{i=1}^n P(A_i) = \alpha, \\
& \sum_{k=1}^n k^2 a(k) = \sum_{i,j=1}^n P(A_i \cap A_j) = \beta, \\
& a(k) \geqslant 0, k = 1, \ldots, n.
\end{aligned}
\tag{2.21}
$$

Then the DS bonud takes the form

$$
\ell_{DS} = \frac{\beta - \lfloor \frac{\beta}{\alpha} \rfloor \alpha}{\lfloor \frac{\beta}{\alpha} \rfloor + 1} + \frac{(\lfloor \frac{\beta}{\alpha} \rfloor + 1)\alpha - \beta}{\lfloor \frac{\beta}{\alpha} \rfloor}. \tag{2.22}
$$

Proof. There are two equality and n inequality constraints for this LP problem. Hence, by Algorithm 1.2.29, there are at most 2 integers $\widehat{k}_1, \widehat{k}_2$ with $1 \leqslant \widehat{k}_1 < \widehat{k}_2 \leqslant n$

such that $a(k) = 0$ for $k \neq \widehat{k}_1, \widehat{k}_2$. Here, we drop the hat notation for simplicity. Hence, the problem simplifies to

$$\begin{aligned}
&\min_{k_1,k_2} a\left(k_1\right) + a\left(k_2\right) \\
&\text{s.t. } k_1 a\left(k_1\right) + k_2 a\left(k_2\right) = \alpha, \\
&\qquad k_1^2 a\left(k_1\right) + k_2^2 a\left(k_2\right) = \beta, \\
&\qquad a\left(k_1\right) \geqslant 0, a\left(k_2\right) \geqslant 0
\end{aligned} \tag{2.23}$$

Solving the 2×2 linear system, we have

$$a(k_1) = \frac{k_2\alpha - \beta}{k_1(k_2 - k_1)},\ a(k_2) = \frac{\beta - k_1\alpha}{k_2(k_2 - k_1)}. \tag{2.24}$$

It was shown in [9] that the optimal choice for k_1 and k_2 is given by $k_1 = \lfloor \frac{\beta}{\alpha} \rfloor$ and $k_2 = \lfloor \frac{\beta}{\alpha} \rfloor + 1$. Substituting the two integers, we get the DS bound

$$\ell_{DS} = \frac{\beta - \lfloor \frac{\beta}{\alpha} \rfloor \alpha}{\lfloor \frac{\beta}{\alpha} \rfloor + 1} + \frac{(\lfloor \frac{\beta}{\alpha} \rfloor + 1)\alpha - \beta}{\lfloor \frac{\beta}{\alpha} \rfloor}. \tag{2.25}$$

□

Remark 2.1.6. *The DS bound originally derived and commonly used in literature [9, 5] takes a different form*

$$\ell_{DS} = \frac{\kappa\theta_1^2}{(2-\kappa)\theta_1 + 2\theta_2} + \frac{(1-\kappa)\theta_1^2}{(1-\kappa)\theta_1 + 2\theta_2}. \tag{2.26}$$

where $\kappa = \frac{2\theta_2}{\theta_1} - \left\lfloor \frac{2\theta_2}{\theta_1} \right\rfloor$. The partial information is given instead as $\theta_1 := \sum_{i=1}^n P\left(A_i\right)$ and $\theta_2 := \sum_{1 \leqslant i < j \leqslant n} P\left(A_i \cap A_j\right)$. We can show that these two expressions are equivalent. Note that $\sum_{i,j} P\left(A_i \cap A_j\right) = 2\sum_{i<j} P\left(A_i \cap A_j\right) + \sum_{i=1}^n P\left(A_i\right)$. Then set $\alpha = \theta_1, \beta = \theta_1 + 2\theta_2, \kappa = \frac{2\theta_2}{\theta_1} - \left\lfloor \frac{2\theta_2}{\theta_1} \right\rfloor$. Hence, $\lfloor \frac{\beta}{\alpha} \rfloor = \lfloor \frac{\theta_1 + 2\theta_2}{\theta_1} \rfloor = 1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor$. We have

$$\begin{aligned}
\ell_{DS} &= \frac{\theta_1 + 2\theta_2 - (1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1}{2 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor} + \frac{(2 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1 - \theta_1 - 2\theta_2}{1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor} \\
&= \frac{(2\theta_2 - \theta_1 \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1}{(2 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1} + \frac{((1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1 - 2\theta_2)\theta_1}{(1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1} \\
&= \frac{(\frac{2\theta_2}{\theta_1} - \lfloor \frac{2\theta_2}{\theta_1} \rfloor)\theta_1^2}{(2 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor - \frac{2\theta_2}{\theta_1})\theta_1 + 2\theta_2} + \frac{((1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor) - \frac{2\theta_2}{\theta_1})\theta_1^2}{(1 + \lfloor \frac{2\theta_2}{\theta_1} \rfloor - \frac{2\theta_2}{\theta_1})\theta_1 + 2\theta_2} \\
&= \frac{\kappa\theta_1^2}{(2-\kappa)\theta_1 + 2\theta_2} + \frac{(1-\kappa)\theta_1^2}{(1-\kappa)\theta_1 + 2\theta_2}.
\end{aligned} \tag{2.27}$$

2.1.2.2 KAT bound

One improvement that can be made to DS bound is replacing the aggregated constraints with the corresponding disaggregated constraints. Then we obtain the Kuai-Alajaji-Takahara (KAT) bound [11].

Proposition 2.1.7. *(KAT bound) The KAT bound is the solution to the following LP problem*

$$\begin{aligned}
\ell_{KAT} := \min_{\{a_i(k), i=1,\dots,n, k=1,\dots,n\}} & \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{a_i(k)}{k} \\
s.t. \ & \sum_{k=1}^{n} a_i(k) = \alpha_i, \quad i = 1,\dots,n, \\
& \sum_{k=1}^{n} k a_i(k) = \beta_i, \quad i = 1,\dots,n, \\
& a_i(k) \geqslant 0, \quad i = 1,\dots,n, \quad k = 1,\dots,n,
\end{aligned} \tag{2.28}$$

Then

$$\ell_{KAT} = \sum_{i=1}^{n} \left\{ \left[\frac{1}{\left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor} - \frac{\frac{\beta_i}{\alpha_i} - \left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor}{\left(1 + \left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor\right)\left(\left\lfloor \frac{\beta_i}{\alpha_i} \right\rfloor\right)} \right] \alpha_i \right\}. \tag{2.29}$$

The derivation for this bound can be found in the reference [11].

Remark 2.1.8. *The KAT bound is sharper than the DS bound and the DC bound by a factor of at most $\frac{9}{8}$, i.e.* $\max\{\ell_{DS}, \ell_{DC}\} \leqslant \ell_{KAT} \leqslant \frac{9}{8}\ell_{DC}$ *[5].*

2.1.2.3 YAT bound

While the KAT bound improves on the existing DS bound and DC bound, it is not the optimal bound given those disaggregated constraints. In short, ℓ_{KAT} is the largest lower bound for $\sum_{i=1}^{n}\sum_{k=1}^{n}\frac{a_i(k)}{k}$ given the disaggregated constraints but it is not the largest lower bound for $\mathbb{P}(\mu \geqslant 1)$. By Lemma 1.2.8, a bound is optimal if we can always construct a sequence of events that attains the bound. It is shown in reference [12] that this is not always possible for the KAT bound.

For the optimal lower bound for $\mathbb{P}(\mu \geqslant 1)$ given constraints $\sum_{k=1}^{n} a_i(k) = \alpha_i$, $\sum_{k=1}^{n} k a_i(k) = \beta_i$ with $i = 1,\dots,n$, Yang, Alajaji, Takahara [12] considered the following LP

$$
\begin{aligned}
\min_{\{a_i(k), i=1,\ldots,n, k=1,\ldots,n\}} & \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{a_i(k)}{k} \\
\text{s.t. } & \sum_{k=1}^{n} a_i(k) = \alpha_i, \quad i = 1, \ldots, n, \\
& \sum_{k=1}^{n} k a_i(k) = \beta_i, \quad i = 1, \ldots, n, \\
& \sum_{i=1}^{n} a_i(k) \geqslant k a_j(k), \quad j, k = 1, \ldots, n, \\
& a_i(k) \geqslant 0, \quad i = 1, \ldots, n, \quad k = 1, \ldots, n.
\end{aligned}
\tag{2.30}
$$

and showed that by adding the constraints $\sum_{i=1}^{n} a_i(k) \geqslant k a_j(k)$ for $j, k = 1, \ldots, n$, this LP produces the optimal lower bound.

However, it is very challenging to derive an analytical bound from the corresponding LP with these extra constraints. Yang, Alajaji, Takahara [12] proposed a relaxed LP by removing the constraints $\sum_{i=1}^{n} a_i(k) \geqslant k a_j(k)$ for $j = 1, \ldots, n$ and $k = 1, \ldots, n-1$, leaving only the constraints $\sum_{i=1}^{n} a_i(n) \geqslant n a_j(n)$ for $j = 1, \ldots, n$. By a relaxed LP problem, we mean an LP with certain constraints removed from the construction. Note that these constraints are equivalent to $a_1(n) = \cdots = a_n(n) \geqslant 0$. To show this, suppose for a contradiction that $\sum_{i=1}^{n} a_i(n) > n a_j(n)$ for some $j = j'$. Then $n \sum_{i=1}^{n} a_i(n) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_i(n) > \sum_{j=1}^{n} n a_j(n) = n \sum_{i=1}^{n} a_i(n)$, which is a contradiction.

Intuitively, the constraints state that degree n event contained in A_i has the same probability for $i = 1, \ldots, n$. This is correct since a degree n event is an event that all n events occur so it is contained in all A_i's. These additional constraints $a_1(n) = \cdots = a_n(n) = x \geqslant 0$ improve the existing KAT bound and the details are given in the following. We will call this bound the Yang-Alajaji-Takahara (YAT) bound [12].

Proposition 2.1.9. *(YAT bound) Define the YAT bound as the solution of the re-*

laxation of the LP problem (2.30) given as follows

$$
\begin{aligned}
\ell_{YAT} := & \min_{\{a_i(k), i=1,...,n, k=1,...,n\}} \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{a_i(k)}{k} \\
& s.t. \ \sum_{k=1}^{n} a_i(k) = \alpha_i, \ \sum_{k=1}^{n} k a_i(k) = \beta_i, \ i = 1, ..., n, \\
& a_1(n) = \cdots = a_n(n) \geqslant 0, \\
& a_i(k) \geqslant 0, \ i, k = 1, ..., n.
\end{aligned} \tag{2.31}
$$

Then YAT bound is given by

$$
P\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \ell_{YAT} = \delta + \sum_{i=1}^{n} \left\{ \left[\frac{1}{\chi\left(\frac{\beta_i'}{\alpha_i'}\right)} - \frac{\frac{\beta_i'}{\alpha_i'} - \chi\left(\frac{\beta_i'}{\alpha_i'}\right)}{\left[1+\chi\left(\frac{\beta_i'}{\alpha_i'}\right)\right]\left[\chi\left(\frac{\beta_i'}{\alpha_i'}\right)\right]} \right] \alpha_i' \right\}. \tag{2.32}
$$

where the function $\chi(\cdot)$ is defined by

$$
\chi(x) = \begin{cases} n-1 & \textit{if } x = n \textit{ where } n \geqslant 2 \textit{ is a integer} \\ \lfloor x \rfloor & \textit{otherwise} \end{cases} \tag{2.33}
$$

and

$$
\delta := \left\{ \max_i \left[\beta_i - (n-1)\alpha_i \right] \right\}^+ \geqslant 0, \quad \alpha_i' := \alpha_i - \delta, \quad \beta_i' := \beta_i - n\delta. \tag{2.34}
$$

The proof for this bound is very extensive and can be found in [12].

Remark 2.1.10. *In the work by Yang et al. [12], it is shown that the YAT bound is at least as good as the KAT bound. This is true by construction since the LP in the YAT bound imposes more constraints, making the feasible set smaller. Hence, the minimum of the objective function is either unchanged if the minimiser is contained in the new feasible set or larger if the minimiser is excluded. This also shows that this particular relaxation of the LP problem that gives the optimal numerical lower bound is an effective extra condition that can be imposed on more analytical bounds derived from LP problems.*

Remark 2.1.11. *(Relationships between different bounds) We will now give an overview for the bounds that have been discussed. The Cauchy-Schwarz type bound ℓ_{CS1} and the DS bound ℓ_{DS} are bounds derived from the aggregated partial information S_1 and S_2. The DS bound is sharper than the Cauchy-Schwarz type bound as it is constructed using an LP. Similarly, the DC bound ℓ_{DC} and the KAT bound ℓ_{KAT} are the corresponding bounds derived from the disaggregated partial information S_1^i and S_2^i for*

$i = 1, \ldots, n$. The KAT bound is sharper than the DC type bound as it is constructed using an LP. While the DS bound is optimal, the KAT bound is not. By imposing the additional constraints $a_1(n) = \cdots = a_n(n) = x \geqslant 0$, the YAT bound ℓ_{YAT} is sharper than the KAT bound by construction. The Prékopa-Gao bound mentioned in [2] includes the additional partial information S_3^i for $i = 1, \ldots, n$. It is, however, not necessarily sharper than the YAT bound [12].

2.2 Analytical lower bounds for $\mathbb{P}(\mu \geqslant m)$

The research on lower bounds for $\mathbb{P}(\mu \geqslant m)$ for a general m is relatively more limited. The work by Boros and Prékopa [10] gave the closed-form bound for $\mathrm{P}(\mu \geqslant m)$ using the aggregated partial information $S_1, \ldots, S_t$ where $t = 2, 3, 4$. Prékopa and Gao [2] generalised the work [10] and derived the closed-form bound for $\mathrm{P}(\mu \geqslant 1)$ using the corresponding disaggregated information $S_1^i, \ldots, S_t^i$ where $t = 2, 3, 4$ and $i = 1, \ldots, n$.

However, there are limited attempts to derive the lower bounds for $\mathrm{P}(\mu \geqslant m)$ for a general m using the disaggregated information $S_1^i, \ldots, S_t^i$ where $i = 1, \ldots, n$. We aim to address this problem for the case $(m, t) = (2, 2)$ and $(3, 3)$.

Chapter 3

New Analytical Lower Bounds for $\mathbb{P}(\mu \geqslant 2)$

We will start with the simpler case where $m = 2$ and $t = 2$. As shown in Lemma 1.2.19, the partial information S_1^i, S_2^i, $i = 1, \ldots, n$ is equivalent to the partial information $\sum_{k=1}^{n} a_i(k), \sum_{k=1}^{n} k a_i(k)$ for $i = 1, \ldots, n$. We will construct the bound finding problem for $\mathbb{P}(\mu \geqslant 2)$ in terms of these two classes of partial information.

Inspired by the additional constraints used in the construction of the YAT bound [12], $a_1(n) = \cdots = a_n(n) \geqslant 0$, we will always impose these additional constraints on our bound finding problems.

In order to avoid trivial results, We will assume that $n \geqslant 4$ for our proof to work. In practice, n is usually reasonably large with $n \gg t$ so this assumption would not be restrictive.

3.1 Bounds through the Cauchy-Schwarz inequality

We are concerned about finding a lower bound $\ell(\theta)$ with disaggregated partial information

$$\theta = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) \tag{3.1}$$

such that

$$\begin{aligned}
\ell(\theta) \leqslant \mathbb{P}(\mu \geqslant 2) = \sum_{i=1}^{n}\sum_{k=2}^{n} \frac{a_i(k)}{k} \\
s.t. \sum_{k=1}^{n} a_i(k) = \alpha_i,\ i = 1, ..., n, \\
\sum_{k=1}^{n} k a_i(k) = \beta_i,\ i = 1, ..., n, \\
a_1(n) = ... = a_n(n) \geqslant 0, \\
a_i(k) \geqslant 0,\ i, k = 1, ..., n.
\end{aligned} \tag{3.2}$$

Denote $x := a_1(n) = ... = a_n(n)$. We can apply the Cauchy-Schwarz inequality on the objective function and derive the following result.

Theorem 3.1.1. *(CS bound for $\mathbb{P}(\mu \geqslant 2)$) For $n \geqslant 4$, the CS type bound $\ell_{2,CS}$ for $\mathbb{P}(\mu \geqslant 2)$ is given in the following*

$$\mathbb{P}(\mu \geqslant 2) \geqslant \ell_{2,CS} := \sum_{i=1}^{n} \left(\frac{\beta_i - \alpha_i}{(n-1)(n-2)} - \frac{2\delta}{n(n-2)} \right) \tag{3.3}$$

where

$$\delta := \min_{i=1,\ldots,n} \left\{ \frac{\beta_i - \alpha_i}{n-1} \right\}. \tag{3.4}$$

Proof. We will proceed the proof in the following four steps:

1. Fix variable $x \geqslant 0$ and consider n sub-problems for each $i = 1, \ldots, n$.
2. Fix variable $a_i(1) \geqslant 0$ and rewrite the problem in suitable forms.
3. Apply the Cauchy-Schwarz inequality.
4. Solve the resulting optimisation problem with decision variables $a_i(1)$ and x.

Step 1 Fixing $x \geqslant 0$ and simplify, we aim to find an ℓ such that

$$\begin{aligned}
\ell \leqslant \sum_{i=1}^{n}\sum_{k=2}^{n-1} \frac{a_i(k)}{k} + x \\
s.t. \sum_{k=1}^{n-1} a_i(k) = \alpha_i - x,\ i = 1, ..., n, \\
\sum_{k=1}^{n-1} k a_i(k) = \beta_i - nx,\ i = 1, ..., n, \\
a_i(k) \geqslant 0,\ i = 1, ..., n, k = 1, ..., n-1.
\end{aligned} \tag{3.5}$$

Note that in (3.5) the constraints and the objective function are separable in i. Then we have the following family of problems of finding lower bounds ℓ_i for each $i = 1, ..., n$ and combine them using $\ell = \sum_{i=1}^{n} \ell_i$:

$$\begin{aligned} \ell_i \leqslant & \sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} \\ s.t. & \sum_{k=1}^{n-1} a_i(k) = \alpha_i - x, \\ & \sum_{k=1}^{n-1} k a_i(k) = \beta_i - nx, \\ & a_i(k) \geqslant 0, \ k = 1, ..., n-1. \end{aligned} \tag{3.6}$$

Step 2 Fix $a_i(1) =: u_i \geqslant 0$ and rearrange:

$$\begin{aligned} \ell_i \leqslant & \sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} \\ s.t. & \sum_{k=2}^{n-1} a_i(k) = \alpha_i - x - u_i, \\ & \sum_{k=2}^{n-1} k a_i(k) = \beta_i - nx - u_i, \\ & a_i(k) \geqslant 0, \ k = 2, ..., n-1. \end{aligned} \tag{3.7}$$

To avoid trivial solutions, here we assume that $\alpha_i - x - u_i, \beta_i - nx - u_i > 0$.

Step 3 By the Cauchy-Schwarz inequality, we have

$$\left(\sum_{k=2}^{n-1} a_i(k) \right)^2 = \left(\sum_{k=2}^{n-1} \sqrt{\frac{a_i(k)}{k}} \sqrt{k a_i(k)} \right)^2 \tag{3.8}$$

$$\leqslant \left(\sum_{k=2}^{n-1} \frac{a_i(k)}{k} \right) \left(\sum_{k=2}^{n-1} k a_i(k) \right). \tag{3.9}$$

Rearranging, we have

$$\sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} \geqslant \frac{\left(\sum_{k=2}^{n-1} a_i(k)\right)^2}{\sum_{k=2}^{n-1} k a_i(k)} + \frac{x}{n} \tag{3.10}$$

$$= \frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \tag{3.11}$$

$$\geqslant \min_{u_i,x} \left(\frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \right). \tag{3.12}$$

Hence, we have the following LP problem and can define its solution as the CS type bound for the $m = 2$ case.

$$\begin{aligned} \ell_{2,\mathrm{CS}} := & \min_{u_i,x,a_i(k):k=2,\ldots,n-1} \left(\frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \right) \\ s.t. \ & \sum_{k=2}^{n-1} a_i(k) = \alpha_i - x - u_i, \\ & \sum_{k=2}^{n-1} k a_i(k) = \beta_i - nx - u_i, \\ & a_i(k) \geqslant 0, \ k = 2, ..., n-1. \end{aligned} \tag{3.13}$$

Step 4 First we fix $x, a_i(k) : k = 1, \ldots, n$ and minimise the objective function over u_i. We can calculate its partial derivative w.r.t u_i.

$$\begin{aligned} \frac{\partial}{\partial u_i} \left(\frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \right) &= \frac{-2(\beta_i - nx - u_i)(\alpha_i - x - u_i) + (\alpha_i - x - u_i)^2}{(\beta_i - nx - u_i)^2} \\ &= \frac{(\alpha_i - x - u_i)((\alpha_i - x - u_i) - 2(\beta_i - nx - u_i))}{(\beta_i - nx - u_i)^2} \\ &< 0 \end{aligned} \tag{3.14}$$

The last inequality holds since

$$2\sum_{k=2}^{n-1} a_i(k) \leqslant \sum_{k=2}^{n-1} k a_i(k) \implies 2(\alpha_i - x - u_i) \leqslant \beta_i - nx - u_i \tag{3.15}$$

$$\implies \alpha_i - x - u_i < 2(\beta_i - nx - u_i) \tag{3.16}$$

when $\alpha_i - x - u_i, \beta_i - nx - u_i \neq 0$ by assumption.

Hence, the objective function is decreasing in u_i so we need to maximise u_i to minimise it. This implies that $u_i > 0$ at the minimum of the objective function.

Now we fix $x \geqslant 0$ instead. Consider the LP problem as follows

$$\begin{aligned} f_i^{2,\mathrm{CS}}(x) := & \min_{u_i, a_i(k): k=2,\ldots,n-1} \left(\frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \right) \\ & s.t. \sum_{k=2}^{n-1} a_i(k) = \alpha_i - x - u_i, \\ & \qquad \sum_{k=2}^{n-1} k a_i(k) = \beta_i - nx - u_i, \\ & \qquad a_i(k) \geqslant 0,\ k = 2, \ldots, n-1, \\ & \qquad u_i > 0. \end{aligned} \tag{3.17}$$

Now we have $(n-1)$ decision variables with 2 equality constraints and $(n-1)$ inequality constraints. We know the constraint $u_i > 0$ is inactive so we need $(n-3)$ out of the rest $(n-2)$ constraints to be active. Then there exists an integer k_1 with $2 \leqslant k_1 \leqslant n-1$ such that $a_i(k) = 0$ for $k \neq k_1$. Then we have

$$a_i(k_1) = \alpha_i - x - u_i \tag{3.18}$$

$$k_1 a_i(k_1) = \beta_i - nx - u_i. \tag{3.19}$$

Solve the linear system, we get

$$u_i = \alpha_i - x - a_i(k_1) \tag{3.20}$$

$$a_i(k_1) = \frac{(\beta_i - nx) - (\alpha_i - x)}{k_1 - 1}. \tag{3.21}$$

Note that u_i is decreasing in $a_i(k_1)$. Hence,

$$\text{maximise } u_i \iff \text{minimise } a_i(k_1). \tag{3.22}$$

Note that $\beta_i - nx = \sum_{k=1}^{n-1} k a_i(k) \geqslant \sum_{k=1}^{n-1} a_i(k) = \alpha_i - x$ so $a_i(k_1)$ is decreasing in k_1. Choosing the largest k_1, we set $k_1 = n-1$. Then

$$a_i(n-1) = \frac{(\beta_i - nx) - (\alpha_i - x)}{n-2} \geqslant 0. \tag{3.23}$$

This gives a condition on x that

$$x \leqslant \frac{\beta_i - \alpha_i}{n-1} \text{ for } i = 1, \ldots, n \tag{3.24}$$

$$\iff x \leqslant \delta \text{ where } \delta := \min_{i=1,\ldots,n} \left\{ \frac{\beta_i - \alpha_i}{n-1} \right\}. \tag{3.25}$$

Substituting all expressions into the objective function, we obtain the following result.

$$
\begin{aligned}
f_i^{2,\mathrm{CS}}(x) = & \min_{u_i, a_i(k): k=2,\dots,n-1} \left(\frac{(\alpha_i - x - u_i)^2}{\beta_i - nx - u_i} + \frac{x}{n} \right) && (3.26)\\
= & \frac{a_i(n-1)^2}{(n-1)a_i(n-1)} + \frac{x}{n} && (3.27)\\
= & \frac{a_i(n-1)}{n-1} + \frac{x}{n} && (3.28)\\
= & \frac{(\beta_i - \alpha_i) - (n-1)x}{(n-1)(n-2)} + \frac{x}{n} && (3.29)\\
= & \frac{\beta_i - \alpha_i}{(n-1)(n-2)} - \frac{2x}{n(n-2)} && (3.30)
\end{aligned}
$$

Since $f_i^{2,\mathrm{CS}}(x)$ is decreasing in x, we need to maximise x to minimise the objective function. By condition (3.25) for x, we choose $x = \delta$.

Hence, the Cauchy-Schwarz bound is given by

$$
\ell_{2,\mathrm{CS}} = \sum_{i=1}^{n} f_i^{2,\mathrm{CS}}(\delta) = \sum_{i=1}^{n} \left(\frac{\beta_i - \alpha_i}{(n-1)(n-2)} - \frac{2\delta}{n(n-2)} \right) \tag{3.31}
$$

where

$$
\delta := \min_{i=1,\dots,n} \left\{ \frac{\beta_i - \alpha_i}{n-1} \right\}. \tag{3.32}
$$

□

3.2 Bounds through the construction of LP problems

We can define the corresponding LP bound as the solution to the following LP:

$$
\begin{aligned}
\mathbb{P}(\mu \geqslant 2) \geqslant \ell_{2,LP} := & \min_{a_i(k):\, i,k=1,\dots,n} \sum_{i=1}^{n} \sum_{k=m}^{n} \frac{a_i(k)}{k} \\
& s.t. \sum_{k=1}^{n} a_i(k) = \alpha_i, \ i = 1, \dots, n, \\
& \sum_{k=1}^{n} k a_i(k) = \beta_i, \ i = 1, \dots, n, \\
& a_1(n) = \dots = a_n(n) \geqslant 0, \\
& a_i(k) \geqslant 0, \ i, k = 1, \dots, n.
\end{aligned}
\tag{3.33}
$$

Denote and fix $x := a_1(n) = ... = a_n(n) \geqslant 0$. Considering each problem $i = 1, ..., n$ separately, we get

$$\begin{aligned} f_i^{2,\mathrm{LP}}(x) := & \min_{a_i(k):\, k=1,\ldots,n-1} \sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} \\ & s.t. \sum_{k=1}^{n-1} a_i(k) = \alpha_i - x,\ \sum_{k=1}^{n-1} k a_i(k) = \beta_i - nx, \\ & \quad a_i(k) \geqslant 0,\ k = 1, ..., n-1. \end{aligned} \tag{3.34}$$

We have $(n-1)$ decision variables in total. Note that there are two equality constraints and $(n-1)$ inequality constraints. Then by the Algorithm 1.2.29, we must have $(n-3)$ active inequality constraints at the optimal point. Hence, there are at most two integers $\widehat{k}_1 < \widehat{k}_2$ between 1 to $n-1$ such that $a_i(k) = 0$ for $k \neq \widehat{k}_1, \widehat{k}_2$.

Before going into the details of the derivations, one important result we will use is the invertibility of Vandermonde matrices.

Lemma 3.2.1. *(Invertibility of Vandermonde matrices) Consider the following Vandermonde matrix:*

$$V[x_1, \ldots, x_n] = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{bmatrix}. \tag{3.35}$$

Then $V[x_1, \ldots, x_n]$ is invertible iff $x_1, \ldots, x_n$ are distinct [17].

The proof for this lemma can be found in [17]. Now we are ready to prove the optimality conditions of the LP problem we are trying to solve.

Lemma 3.2.2. *(Optimality conditions) For $n \geqslant 4$, the minimum of (3.33) is achieved at $\widehat{k}_1 = 1$.*

Proof. To show this, note that we can have at most two positive variables $a_i(\widehat{k}_1)$, $a_i(\widehat{k}_2)$ as $a_i(k) = 0$ for $k \neq \widehat{k}_1, \widehat{k}_2$. In addition, $a_i(k)$ contributes to the objective function $\sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n}$ only when $k \geqslant 2$, i.e. $a_i(1)$ does not contribute. Hence, we can consider the following three variables $a_i(1), a_i(k_1), a_i(k_2)$ with $2 \leqslant k_1 < k_2 \leqslant n-1$ and $a_i(k) = 0$ for $k \neq 1, k_1, k_2$. We will set one of $a_i(1), a_i(k_1), a_i(k_2)$ to be zero so that there are at most two positive decision variables. Under this construction, only $a_i(k_1), a_i(k_2)$ can contribute to the objective function where the objective function is now given by

$$\sum_{k=2}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} = \frac{a_i(k_1)}{k_1} + \frac{a_i(k_2)}{k_2} + \frac{x}{n}. \tag{3.36}$$

We will prove the result in the following three steps:

1. Assuming $a_i(1)$ are fixed, find $a_i(k_1), a_i(k_2)$, in terms of $a_i(1)$.
2. Find the partial derivatives of the objective function w.r.t $a_i(1)$.
3. Prove that these derivatives are negative and thus show the optimality at $\widehat{k}_1 = 1$.

Step 1 From (3.34), we get the following two equations.

$$a_i(1) + a_i(k_1) + a_i(k_2) = \alpha_i - x \tag{3.37}$$

$$a_i(1) + k_1 a_i(k_1) + k_2 a_i(k_2) = \beta_i - nx \tag{3.38}$$

Moving $a_i(1)$ to the right, we have

$$a_i(k_1) + a_i(k_2) = \alpha_i - x - a_i(1) \tag{3.39}$$

$$k_1 a_i(k_1) + k_2 a_i(k_2) = \beta_i - nx - a_i(1). \tag{3.40}$$

This can be written as a linear system $Ka = b$ with

$$K = \begin{bmatrix} 1 & 1 \\ k_1 & k_2 \end{bmatrix} \qquad a = \begin{bmatrix} a_i(k_1) \\ a_i(k_2) \end{bmatrix} \qquad b = \begin{bmatrix} \alpha_i - x - a_i(1) \\ \beta_i - nx - a_i(1) \end{bmatrix}. \tag{3.41}$$

Note that K is the transpose of a Vandermonde matrix with a second column $(k_1, k_2)^\top$. Since $k_1 < k_2$, then by Lemma 3.2.1, $A^\top$ is invertible and thus A is invertible. Hence, there is a unique solution to this linear system which is given by

$$a_i(k_1) = \frac{k_2(\alpha_i - x - a_i(1)) - (\beta_i - nx - a_i(1))}{k_2 - k_1}, \tag{3.42}$$

$$a_i(k_2) = \frac{(\beta_i - nx - a_i(1)) - k_1(\alpha_i - x - a_i(1))}{k_2 - k_1}. \tag{3.43}$$

Step 2 To calculate the partial derivatives of the objective function w.r.t $a_i(1)$, we can first calculate the derivatives of $a_i(k_1), a_i(k_2)$ w.r.t $a_i(1)$.

$$\frac{\partial a_i(k_1)}{\partial a_i(1)} = \frac{1 - k_2}{k_2 - k_1} \qquad \frac{\partial a_i(k_2)}{\partial a_i(1)} = \frac{k_1 - 1}{k_2 - k_1} \tag{3.44}$$

Then the partial derivatives of the objective function w.r.t $a_i(1)$ is given by

$$\frac{\partial}{\partial a_i(1)}\left(\frac{a_i(k_1)}{k_1}+\frac{a_i(k_2)}{k_2}+\frac{x}{n}\right)=\frac{1-k_2}{k_1(k_2-k_1)}+\frac{k_1-1}{k_2(k_2-k_1)} \tag{3.45}$$

$$=\frac{1}{k_2-k_1}\left(\frac{k_2-k_2^2+k_1^2-k_1}{k_1k_2}\right) \tag{3.46}$$

$$=\frac{1}{k_2-k_1}\left(\frac{(k_2-k_1)(1-(k_2+k_1))}{k_1k_2}\right) \tag{3.47}$$

$$=\frac{1-(k_2+k_1)}{k_1k_2}<0. \tag{3.48}$$

Step 3 The last inequality holds since $2 \leqslant k_1 < k_2$. Intuitively, this shows that objective function is decreasing in $a_i(1)$ so we need to maximise $a_i(1)$ to minimise the objective function. Hence, the minimum of the objective function cannot occur at $a_i(1) = 0$ while $a_i(k_1), a_i(k_2) > 0$ as we can increase $a_i(1)$ to any small positive number to achieve a smaller objective function. Hence, the minimum can only occur when at least one of $a_i(k_1), a_i(k_2)$ vanishes. Equivalently, the minimum of (3.33) is achieved at $\widehat{k}_1 = 1$. $\square$

With this lemma, we are ready to derive our LP bound.

Theorem 3.2.3. *(LP bound for $\mathbb{P}(\mu \geqslant 2)$) Define the LP bound as in (3.33). Then $\ell_{2,LP} = \ell_{2,CS}$.*

Proof. By Lemma 3.2.2 and Equation 3.34 and by dropping the hat on $\widehat{k}_2$, we have the following LP problem.

$$\begin{aligned}\min_{x,a_i(k):\,k=1,k_2} \quad & \frac{a_i(k_2)}{k_2}+\frac{x}{n}\\ s.t.\ & a_i(1)+a_i(k_2)=\alpha_i-x,\\ & a_i(1)+k_2a_i(k_2)=\beta_i-nx,\\ & a_i(1),a_i(k_2)\geqslant 0.\end{aligned} \tag{3.49}$$

Solving the system of two equations, we have

$$\begin{aligned}a_i(1)&=\frac{k_2(\alpha_i-x)-(\beta_i-nx)}{k_2-1}\geqslant 0,\\ a_i(k_2)&=\frac{(\beta_i-nx)-(\alpha_i-x)}{k_2-1}\geqslant 0.\end{aligned} \tag{3.50}$$

Note that $\frac{a_i(k_2)}{k_2} + \frac{x}{n}$ is decreasing in k_2. Hence, we need to choose the largest k_2 possible, i.e. $k_2 = n - 1$. Then we have

$$\begin{aligned} f_i^{2,\mathrm{LP}}(x) &= \min_{a_i(k):\, k=1,k_2} \frac{a_i(k_2)}{k_2} + \frac{x}{n} \\ &= \frac{(\beta_i - nx) - (\alpha_i - x)}{(n-1)(n-2)} + \frac{x}{n} \\ &= \frac{\beta_i - \alpha_i}{(n-1)(n-2)} - \frac{2x}{n(n-2)}. \end{aligned} \tag{3.51}$$

Here, $f_i^{2,\mathrm{LP}}(x)$ is decreasing in x so we need to maximise x to minimise $f_i^{2,\mathrm{LP}}(x)$. Note that the two inequalities (3.50) with $k_2 = n - 1$ give the following two constraints on x.

$$\beta_i - (n-1)\alpha_i \leqslant x \leqslant \frac{\beta_i - \alpha_i}{n-1} \tag{3.52}$$

Note that these two inequalities hold for $i = 1, \ldots, n$ and in addition $x \geqslant 0$. Then we have the following constraints on x.

$$\left[\max_{i=1,\ldots,n} \{\beta_i - (n-1)\alpha_i \leqslant x\}\right]^+ \leqslant x \leqslant \min_{i=1,\ldots,n} \left\{\frac{\beta_i - \alpha_i}{n-1}\right\} =: \delta \tag{3.53}$$

where $[a]^+ = \max(a, 0)$. The largest x we can choose is δ. Hence, we get the following bound for $\mathbb{P}(\mu \geqslant 2)$.

$$\ell_{2,\mathrm{LP}} = \sum_{i=1}^{n} f_i^{2,\mathrm{LP}}(\delta) = \sum_{i=1}^{n} \left(\frac{\beta_i - \alpha_i}{(n-1)(n-2)} - \frac{2\delta}{n(n-2)}\right) = \ell_{2,\mathrm{CS}} \tag{3.54}$$

where

$$\delta := \min_{i=1,\ldots,n} \left\{\frac{\beta_i - \alpha_i}{n-1}\right\}. \tag{3.55}$$

Hence, $\ell_{2,\mathrm{LP}} = \ell_{2,\mathrm{CS}}$. □

Remark 3.2.4. *It is a surprising result that the Cauchy-Schwarz type bound and the LP bound are equal in this case. In the literature, the Cauchy-Schwarz type bound is usually very loose and less sharp than the corresponding LP bound. The key difference here is that we applied the additional constraints $a_1(n) = \ldots = a_n(n) = x \geqslant 0$ to the derivation of the Cauchy-Schwarz type bound, which is never seen in the literature.*

To see why the two bounds are equal, let us examine the Cauchy-Schwarz inequality applied at the minimiser of the LP.

$$\left(\sum_{k=2}^{n-1} a_i(k)\right)^2 \leqslant \left(\sum_{k=2}^{n-1} \frac{a_i(k)}{k}\right) \left(\sum_{k=2}^{n-1} k a_i(k)\right) \tag{3.56}$$

At the optimal point, $a_i(k) = 0$ for $2 \leqslant k \leqslant n-2$. Hence, the inequality becomes

$$(a_i(n-1))^2 \leqslant \left(\frac{a_i(n-1)}{n-1}\right)((n-1)a_i(n-1)) = (a_i(n-1))^2 . \tag{3.57}$$

Notice that the inequality holds as an equality in this case. Hence, the Cauchy-Schwarz inequality does not change the minimum of the objective function, i.e. the Cauchy-Schwarz type bound and the LP bound are necessarily equal.

Remark 3.2.5. *Comparing the $\ell_{2,LP}$ bound to the bound by Boros and Prékopa [10] using the aggregated partial information S_1, S_2, $\ell_{2,LP}$ is sharper since it uses the disaggregated partial information equivalent to S_1^i, S_2^i, $i = 1, \ldots, n$. The reasoning is given in Remark 2.1.10. In short, when more constraints are imposed, the feasible set of the minimisation problem is smaller, making the minimum of the objective function larger.*

Chapter 4

New Analytical Lower Bounds for $\mathbb{P}(\mu \geqslant 3)$

In this section, we will continue our discussion of analytical lower bounds for the case where $m = 3$ and $t = 3$, i.e. finding lower bounds for $\mathbb{P}(\mu \geqslant 3)$ using the partial information S_1^i, S_2^i, S_3^i, $i = 1, \ldots, n$. This is equivalent to giving the partial information $\sum_{k=1}^n a_i(k), \sum_{k=1}^n ka_i(k), \sum_{k=1}^n k^2 a_i(k)$ for $i = 1, \ldots, n$ by Lemma 1.2.19.

Note that the Cauchy-Schwarz inequality requires the inner product of two terms. In this context, it means we need two pieces of partial information. However, we are given three different classes of partial information. It becomes unclear how to generalise the use of the Cauchy-Schwarz inequality to the case $t = 3$. Thus, we will focus on the construction of bounds through LPs.

4.1 Bounds through the construction of LP problems

Define the LP bound as the minimum of the following LP problem:

$$\begin{aligned}
\mathbb{P}(\mu \geqslant 3) \geqslant \ell_{3,\mathrm{LP}} := & \min_{a_i(k):\, i,k=1,\dots,n} \sum_{i=1}^{n} \sum_{k=3}^{n} \frac{a_i(k)}{k} \\
s.t.\ & \sum_{k=1}^{n} a_i(k) = \alpha_i,\ i = 1, \dots, n, \\
& \sum_{k=1}^{n} k a_i(k) = \beta_i,\ i = 1, \dots, n, \\
& \sum_{k=1}^{n} k^2 a_i(k) = \gamma_i,\ i = 1, \dots, n, \\
& a_1(n) = \dots = a_n(n) \geqslant 0, \\
& a_i(k) \geqslant 0,\ i, k = 1, \dots, n.
\end{aligned} \tag{4.1}$$

Denote and fix $x := a_1(n) = \dots = a_n(n) \geqslant 0$. Considering each problem $i = 1, \dots, n$ separately, we get

$$\begin{aligned}
f_i^{3,\mathrm{LP}}(x) := & \min_{a_i(k):\, k=1,\dots,n-1} \sum_{k=3}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} \\
s.t.\ & \sum_{k=1}^{n-1} a_i(k) = \alpha_i - x, \\
& \sum_{k=1}^{n-1} k a_i(k) = \beta_i - nx, \\
& \sum_{k=1}^{n} k^2 a_i(k) = \gamma_i - n^2 x, \\
& a_i(k) \geqslant 0,\ k = 1, \dots, n-1.
\end{aligned} \tag{4.2}$$

Note that there are three equality constraints and $(n-1)$ inequality constraints. Then by the Algorithm 1.2.29, we must have $(n-4)$ active inequality constraints at the optimal point. Hence, there are at most three integers $\widehat{k}_1, \widehat{k}_2, \widehat{k}_3$ with $1 \leqslant \widehat{k}_1 < \widehat{k}_2 < \widehat{k}_3 \leqslant n-1$ such that $a_i(k) = 0$ for $k \neq \widehat{k}_1, \widehat{k}_2, \widehat{k}_3$.

In order to prove the optimality conditions, we need the the theory of cyclic polynomials to simplify our analysis. We will introduce the definitions and results in the following.

Definition 4.1.1. *(Cyclic polynomials) Let $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ be a polynomial over $\mathbb{R}$ of n variables $x_1, \ldots, x_n$. Let $G \subseteq GL(n, \mathbb{R})$ be a finite matrix group. Then a polynomial $f(\mathbf{x})$ is invariant under G if $f(\mathbf{x}) = f(A \cdot \mathbf{x})\ \forall A \in G$. In particular, if $f(\mathbf{x})$ is invariant under the cyclic permutation $C : x_1 \mapsto x_2, \ldots, x_{n-1} \mapsto x_n, x_n \mapsto x_1$, i.e. $f(x_1, x_2, x_3, \ldots, x_n) = f(x_2, x_3, \ldots, x_n, x_1)$, we call $f(\mathbf{x})$ a cyclic polynomial [18].*

Example 4.1.2. *Consider the polynomial $f(x, y, z) = x^2 + y^2 + z^2$. Then*

$$f(y, z, x) = y^2 + z^2 + x^2 = x^2 + y^2 + z^2 = f(x, y, z). \tag{4.3}$$

Hence, $f(x, y, z)$ is a cyclic polynomial of three variables by definition.

If we know any root to a cyclic polynomial, we can factor the polynomial easily using cyclicity. To formalise this, we have the following lemma.

Lemma 4.1.3. *(Factorisation of cyclic polynomials) Let $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ be a cyclic polynomial over $\mathbb{R}$ of n variables. If $x_1 = x_2$ is a root of $f(\mathbf{x})$, i.e.*

$$f(x_1, x_1, x_3, x_4, \ldots, x_n) = 0, \tag{4.4}$$

then $f(\mathbf{x})$ has the factor $(x_1 - x_2) \cdots (x_{n-1} - x_n)(x_n - x_1)$.

Proof. Suppose $x_1 = x_2$ is a root of $f(\mathbf{x})$. Then by the Factor Theorem [19, p. 376], $f(\mathbf{x})$ has the factor $(x_1 - x_2)$. Then we can write $f(\mathbf{x}) = (x_1 - x_2)g(\mathbf{x})$ where $g(\mathbf{x})$ is some polynomial of n variables. Let C be the same cyclic permutation in 4.1.1. Since $f(\mathbf{x})$ is cyclic, then

$$f(\mathbf{x}) = f(C \cdot \mathbf{x}) = (x_2 - x_3)g(C \cdot \mathbf{x}). \tag{4.5}$$

Since $g(C \cdot \mathbf{x})$ is a polynomial, then $f(\mathbf{x})$ contains a factor $(x_2 - x_3)$. Repeatedly applying the cyclic permutation C, we can show that $(x_1 - x_2), \ldots, (x_{n-1} - x_n), (x_n - x_1)$ are all factors of $f(\mathbf{x})$. Hence, $f(\mathbf{x})$ contains the factor $(x_1 - x_2) \cdots (x_{n-1} - x_n)(x_n - x_1)$. $\square$

Now we are ready to prove the optimality conditions. Similar as before, we will assume that n is reasonably large. In particular, $n \geqslant 6$ will suffice.

Lemma 4.1.4. *(Optimality conditions) For $n \geqslant 6$, the minimum of (4.2) is achieved at $\widehat{k}_1 = 1, \widehat{k}_2 = 2$.*

Proof. To show this, note that we can have at most three positive decision variables $a_i(\widehat{k}_1)$, $a_i(\widehat{k}_2)$, $a_i(\widehat{k}_3)$ as $a_i(k) = 0$ for $k \neq \widehat{k}_1, \widehat{k}_2, \widehat{k}_3$. In addition, $a_i(k)$ contributes to the objective function $\sum_{k=3}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n}$ only when $k \geqslant 3$. Hence, we can consider the following five variables $a_i(1), a_i(2), a_i(k_1), a_i(k_2), a_i(k_3)$ with $3 \leqslant k_1 < k_2 < k_3 \leqslant n-1$ and $a_i(k) = 0$ for $k \neq 1, 2, k_1, k_2, k_3$. We will set two of $a_i(1), a_i(2), a_i(k_1), a_i(k_2), a_i(k_3)$ to be zero so that there are at most three positive decision variables. Under this construction, only $a_i(k_1), a_i(k_2), a_i(k_3)$ can contribute to the objective function where the objective function is now given by

$$\sum_{k=3}^{n-1} \frac{a_i(k)}{k} + \frac{x}{n} = \frac{a_i(k_1)}{k_1} + \frac{a_i(k_2)}{k_2} + \frac{a_i(k_3)}{k_3} + \frac{x}{n}. \tag{4.6}$$

We will prove the result in the following three steps:

1. Assuming $a_i(1), a_i(2)$ are fixed, find $a_i(k_1), a_i(k_2), a_i(k_3)$ in terms of $a_i(1), a_i(2)$.
2. Find the partial derivatives of the objective function w.r.t $a_i(1)$ and $a_i(2)$.
3. Prove that these derivatives are negative and thus show the optimality at $\widehat{k}_1 = 1, \widehat{k}_2 = 2$

Step 1. From (4.2), we get the following three equations.

$$a_i(1) + a_i(2) + a_i(k_1) + a_i(k_2) + a_i(k_3) = \alpha_i - x \tag{4.7}$$
$$a_i(1) + 2a_i(2) + k_1 a_i(k_1) + k_2 a_i(k_2) + k_3 a_i(k_3) = \beta_i - nx \tag{4.8}$$
$$a_i(1) + 4a_i(2) + k_1^2 a_i(k_1) + k_2^2 a_i(k_2) + k_3^2 a_i(k_3) = \gamma_i - n^2 x \tag{4.9}$$

Moving $a_i(1)$ and $a_i(2)$ to the right, we get

$$a_i(k_1) + a_i(k_2) + a_i(k_3) = \alpha_i - x - a_i(1) - a_i(2) \tag{4.10}$$
$$k_1 a_i(k_1) + k_2 a_i(k_2) + k_3 a_i(k_3) = \beta_i - nx - a_i(1) - 2a_i(2) \tag{4.11}$$
$$k_1^2 a_i(k_1) + k_2^2 a_i(k_2) + k_3^2 a_i(k_3) = \gamma_i - n^2 x - a_i(1) - 4a_i(2). \tag{4.12}$$

This can be written as a linear system $Ka = b$ with

$$K = \begin{bmatrix} 1 & 1 & 1 \\ k_1 & k_2 & k_3 \\ k_1^2 & k_2^2 & k_3^2 \end{bmatrix} \qquad a = \begin{bmatrix} a_i(k_1) \\ a_i(k_2) \\ a_i(k_3) \end{bmatrix} \qquad b = \begin{bmatrix} \alpha_i - x - a_i(1) - a_i(2) \\ \beta_i - nx - a_i(1) - 2a_i(2) \\ \gamma_i - n^2 x - a_i(1) - 4a_i(2) \end{bmatrix}. \tag{4.13}$$

Note that $K = V[k_1, k_2, k_3]^\top$ which is the transpose of a Vandermonde matrix. Since $k_1 < k_2 < k_3$, then by Lemma 3.2.1, $A^\top$ is invertible and thus A is invertible. Hence, there is a unique solution to this linear system which is given by

$$\begin{aligned} a_i(k_1) &= \frac{k_3k_2(\alpha_i - x) - (k_3+k_2)(\beta_i - nx) + (\gamma_i - n^2x)}{(k_3-k_1)(k_2-k_1)} \\ &\quad - \frac{(k_3-1)(k_2-1)a_i(1)}{(k_3-k_1)(k_2-k_1)} - \frac{(k_3-2)(k_2-2)a_i(2)}{(k_3-k_1)(k_2-k_1)}, \end{aligned} \tag{4.14}$$

$$\begin{aligned} a_i(k_2) &= \frac{k_3k_1(\alpha_i - x) - (k_3+k_1)(\beta_i - nx) + (\gamma_i - n^2x)}{(k_3-k_2)(k_1-k_2)} \\ &\quad - \frac{(k_3-1)(k_1-1)a_i(1)}{(k_3-k_2)(k_1-k_2)} - \frac{(k_3-2)(k_1-2)a_i(2)}{(k_3-k_2)(k_1-k_2)}, \end{aligned} \tag{4.15}$$

$$\begin{aligned} a_i(k_3) &= \frac{k_2k_1(\alpha_i - x) - (k_2+k_1)(\beta_i - nx) + (\gamma_i - n^2x)}{(k_2-k_3)(k_1-k_3)} \\ &\quad - \frac{(k_2-1)(k_1-1)a_i(1)}{(k_2-k_3)(k_1-k_3)} - \frac{(k_2-2)(k_1-2)a_i(2)}{(k_2-k_3)(k_1-k_3)}. \end{aligned} \tag{4.16}$$

Step 2. To calculate the partial derivatives of the objective function w.r.t $a_i(1)$ and $a_i(2)$, we can first calculate the derivatives of $a_i(k_1), a_i(k_2), a_i(k_3)$ w.r.t $a_i(1)$ and $a_i(2)$.

$$\begin{aligned} \frac{\partial a_i(k_1)}{\partial a_i(1)} &= -\frac{(k_3-1)(k_2-1)}{(k_3-k_1)(k_2-k_1)} & \frac{\partial a_i(k_1)}{\partial a_i(2)} &= -\frac{(k_3-2)(k_2-2)}{(k_3-k_1)(k_2-k_1)} \\ \frac{\partial a_i(k_2)}{\partial a_i(1)} &= -\frac{(k_3-1)(k_1-1)}{(k_3-k_2)(k_1-k_2)} & \frac{\partial a_i(k_2)}{\partial a_i(2)} &= -\frac{(k_3-2)(k_1-2)}{(k_3-k_2)(k_1-k_2)} \\ \frac{\partial a_i(k_3)}{\partial a_i(1)} &= -\frac{(k_2-1)(k_1-1)}{(k_2-k_3)(k_1-k_3)} & \frac{\partial a_i(k_3)}{\partial a_i(2)} &= -\frac{(k_2-2)(k_1-2)}{(k_2-k_3)(k_1-k_3)} \end{aligned} \tag{4.17}$$

Hence, we can obtain the following expressions for the partial derivatives of the objective function w.r.t $a_i(1)$ and $a_i(2)$:

$$\begin{aligned} G_{i1}(k_1,k_2,k_3) &:= \frac{\partial}{\partial a_i(1)}\left(\frac{a_i(k_1)}{k_1} + \frac{a_i(k_2)}{k_2} + \frac{a_i(k_3)}{k_3} + \frac{x}{n}\right) \\ = -\frac{(k_3-1)(k_2-1)}{k_1(k_3-k_1)(k_2-k_1)} &- \frac{(k_3-1)(k_1-1)}{k_2(k_3-k_2)(k_1-k_2)} - \frac{(k_2-1)(k_1-1)}{k_3(k_2-k_3)(k_1-k_3)}, \end{aligned} \tag{4.18}$$

and

$$G_{i2}(k_1,k_2,k_3) := \frac{\partial}{\partial a_i(2)}\left(\frac{a_i(k_1)}{k_1} + \frac{a_i(k_2)}{k_2} + \frac{a_i(k_3)}{k_3} + \frac{x}{n}\right) \tag{4.19}$$

$$= -\frac{(k_3-2)(k_2-2)}{k_1(k_3-k_1)(k_2-k_1)} - \frac{(k_3-2)(k_1-2)}{k_2(k_3-k_2)(k_1-k_2)} - \frac{(k_2-2)(k_1-2)}{k_3(k_2-k_3)(k_1-k_3)}. \tag{4.20}$$

Step 3 Note that the $m = 3$ case is much more challenging to deal with than the $m = 2$ case and it is unclear whether both $G_{i1}(k_1,k_2,k_3), G_{i2}(k_1,k_2,k_3) < 0$. We need a more systematic method to determine the signs of the two functions.

First observe that both $G_{i1}(k_1, k_2, k_3), G_{i2}(k_1, k_2, k_3)$ are sums of three fractions that are very similar. In particular, they are invariant under cyclic permutations, i.e. $G_{i1}(k_1, k_2, k_3) = G_{i1}(k_3, k_1, k_2) = G_{i1}(k_2, k_3, k_1)$ and the same for $G_{i2}(k_1, k_2, k_3)$. The theory of cyclic polynomials can be used to determine the sign.

To simplify the notation, we will use $x, y, z \in \mathbb{R}$ with $3 \leqslant x \leqslant y - 1 \leqslant z - 2$ as the inputs for these two functions and only require x, y, z to be real numbers. Note the x here is a generic real number and not to be confused with $a_i(n) = x, i = 1, \ldots, n$. First consider $G_{i1}(x, y, z)$ and combine the three fractions.

$$\begin{aligned} G_{i1}(x,y,z) &= \frac{-z^3y^2 + z^2y^3 + 2z^3y - 4z^2y - 2zy^3 + 4zy^2}{zyx(z-x)(y-x)(z-y)} \\ &+ \frac{z^3x^2 - z^2x^3 - 2z^3x + 4z^2x + 2zx^3 - 4zx^2}{zyx(z-x)(y-x)(z-y)} \\ &+ \frac{-y^3x^2 + y^2x^3 + 2y^3x - 4y^2x - 2yx^3 + 4yx^2}{zyx(z-x)(y-x)(z-y)} \\ &=: \frac{P_{i1}(x,y,z)}{zyx(z-x)(y-x)(z-y)} \end{aligned} \tag{4.21}$$

where we have defined the numerator as $P_{i1}(x, y, z)$. Note that $P_{i1}(x, y, z)$ is a cyclic polynomial by construction since the denominator is cyclic.

Note that by assumption $x < y < z$. Then the denominator $zyx(z-x)(y-x)(z-y)$ is positive. Hence, we can just work with the numerator $P_{i1}(x, y, z)$ to determine its sign. To determine the sign of a polynomial, one can consider its roots and the corresponding intervals between different roots.

In order to simplify to analysis, we can work with a univariate polynomial. Now fix y and z and consider x as a variable. Then we can treat $P_{i1}(x, y, z)$ as a function of $x \in [3, y-1]$ only, denoted by $F_{i1}(x)$. Then, $F_{i1}(x)$ is automatically cyclic. After grouping different orders of x, we get

$$\begin{aligned} F_{i1}(x) &= (z - z^2 + y^2 - y)x^3 + (z^3 - z - y^3 + y)x^2 + (-z^3 + z^2 + y^3 - y^2)x \\ &+ (-z^3y^2 + z^3y - z^2y + z^2y^3 - zy^3 + zy^2). \end{aligned} \tag{4.22}$$

Consider $F_{i1}(y)$. We have

$$\begin{aligned} F_{i1}(y) &= (z - z^2 + y^2 - y)y^3 + (z^3 - z - y^3 + y)y^2 + (-z^3 + z^2 + y^3 - y^2)y \\ &+ (-z^3y^2 + z^3y - z^2y + z^2y^3 - zy^3 + zy^2) \\ &= (zy^3 - z^2y^3 + y^5 - y^4) + (z^3y^2 - zy^2 - y^5 + y^3) \\ &+ (-z^3y + z^2y + y^4 - y^3) + (-z^3y^2 + z^3y - z^2y + z^2y^3 - zy^3 + zy^2) \\ &= 0. \end{aligned} \tag{4.23}$$

Hence, $x = y$ is a also root of $P_{i1}(x, y, z)$. By Lemma 4.1.3, $P_{i1}(x, y, z)$ has the factor $(x - y)(x - z)(z - y)$ and so does $F_{i1}(x)$. We may factor $F_{i1}(x)$ as follows

$$F_{i1}(x) = (x - y)(x - z)(z - y)((1 - (z + y))x + (z + y - zy - 1)). \tag{4.24}$$

Hence, there are three roots to $F_{i1}(x) = 0$, given by

$$x_1 = y, \tag{4.25}$$

$$x_2 = z, \tag{4.26}$$

$$x_3 = \frac{z + y - zy - 1}{(z + y) - 1}. \tag{4.27}$$

By assumption, $y, z > 2 > 0$ so $x_1, x_2 > 0$. It remains to determine the sign of x_3. Note that $z + y - 1 > 3 > 0$ and

$$z + y - zy - 1 = z(1 - y) + y - 1 < -z + y - 1 < -1 < 0. \tag{4.28}$$

Hence, $x_3 < 0$. Now we have shown that $F_{i1}(x)$ is a cubic polynomial with two positive roots on $x \in (y - 1, \infty)$ and one negative root. Thus, $F_{i1}(x)$ does not change sign on $[3, y - 1]$, which is the interval of concern.

Note that cubic polynomials with three distinct real roots can only have two shapes, determined by the sign of the coefficient of x^3. In this case, the coefficient is

$$(z - y)((1 - (z + y)). \tag{4.29}$$

Since $z - y > 0$ and $1 - (z + y) < 0$, the coefficient is negative. Hence, we have the following plot for $F_{i1}(x)$.

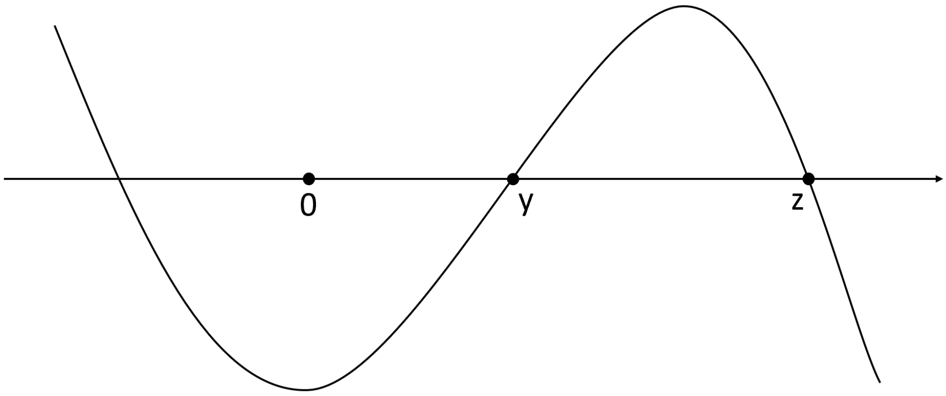

Figure 4.1: $F_{i1}(x)$ plot

This shows that $F_{i1}(x) < 0$ for $x \in [3, y-1]$. Hence, $G_{i1}(x, y, z) < 0$ for $3 \leqslant x < y - 1 \leqslant z - 2$, which implies $G_{i1}(k_1, k_2, k_3) < 0$ for $3 \leqslant k_1 < k_2 < k_3 \leqslant n - 1$.

Similar for $G_{i2}(x, y, z)$, we can define $P_{i2}(x, y, z)$ as its numerator.

$$G_{i2}(x, y, z) =: \frac{P_{i1}(x, y, z)}{zyx(z - x)(y - x)(z - y)} \tag{4.30}$$

Fix y and z and treat $P_{i2}(x, y, z)$ as a function of $x \in [3, y - 1]$ only, denoted by $F_{i2}(x)$. Grouping different orders of x, we have

$$\begin{aligned} F_{i2}(x) &= (2z - z^2 + y^2 - 2y)x^3 + (z^3 - 4z - y^3 + 4y)x^2 \\ &\quad + (-2z^3 + 4z^2 + 2y^3 - 4y^2)x \\ &\quad + (-z^3y^2 + 2z^3y - 4z^2y + z^2y^3 - 2zy^3 + 4zy^2). \end{aligned} \tag{4.31}$$

Similarly, $F_{i2}(y) = 0$. Then $(x-y)(x-z)(z-y)$ is a factor of $F_{i2}(x)$ by Lemma 4.1.3. We can factor $F_{i2}(x)$ as follows

$$F_{i2}(x) = (x - y)(x - z)(z - y)((2 - (z + y)x + (2(z + y) - zy - 4)). \tag{4.32}$$

Three roots of $F_{i2}(x)$ are given by

$$x_4 = y, \tag{4.33}$$

$$x_5 = z, \tag{4.34}$$

$$x_6 = \frac{2(z + y) - zy - 4}{(z + y) - 2}. \tag{4.35}$$

As before, $x_4, x_5 > 0$. It remains to determine the sign of x_6. Note $z + y - 2 > 2 > 0$ and

$$2(z + y) - zy - 4 = \frac{1}{2}(4 - z)y + \frac{1}{2}(4 - y)z - 4 < -4 < 0 \tag{4.36}$$

since by assumption $y \geqslant 4, z \geqslant 5$. Hence, $x_6 < 0$. Now we have shown that $F_{i2}(x)$ is a cubic polynomial with two positive roots on $x \in (y-1, \infty)$ and one negative root. Thus, $F_{i2}(x)$ does not change sign on $[3, y-1]$, which is the interval of concern.

The coefficient of x^3 is given by

$$(z-y)((2-(z+y)). \tag{4.37}$$

Since $z-y>0$ and $2-(z+y)<0$, the coefficient is negative and thus $F_{i2}(x)<0$ for $x \in [3, y-1]$ by the previous argument. Hence, $G_{i2}(x,y,z)<0$ for $3 \leqslant x < y-1 \leqslant z-2$, which implies $G_{i2}(k_1,k_2,k_3)<0$ for $3 \leqslant k_1 < k_2 < k_3 \leqslant n-1$.

Now we have shown that $G_{i1}(k_1,k_2,k_3), G_{i2}(k_1,k_2,k_3) < 0$. Similar as before, this shows that objective function is decreasing in $a_i(1), a_i(2)$ so we need to maximise $a_i(1), a_i(2)$ to minimise the objective function. Hence, the minimum of the objective function cannot occur when either $a_i(1) = 0 = a_i(2)$ while $a_i(k_1), a_i(k_2), a_i(k_3) > 0$ or one of $a_i(1), a_i(2)$ is zero while two of $a_i(k_1), a_i(k_2), a_i(k_3)$ are positive as we can increase $a_i(1)$ or $a_i(2)$ to any small positive number to achieve a smaller objective function. Hence, the minimum can only occur when at least two of $a_i(k_1), a_i(k_2), a_i(k_3)$ vanishes. Equivalently, the minimum of (4.1) is achieved at $\widehat{k}_1 = 1, \widehat{k}_2 = 2$. □

Now we are ready to derive the LP bound for $\mathbb{P}(\mu \geqslant 3)$.

Theorem 4.1.5. *(LP bound for 3-out-of-n events) The LP bound $\ell_{3,LP}$ takes the following form:*

$$\mathbb{P}(\mu \geqslant 3) \geqslant \ell_{3,LP} = \sum_{i=1}^{n} \frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)(n-3)} - \frac{3\delta}{n(n-3)} \tag{4.38}$$

where

$$\delta := \min\left\{ \min_{i=1,\dots,n} \left\{\frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)}\right\}, \min_{i=1,\dots,n} \left\{\frac{2(n-1)\alpha_i - (n+1)\beta_i + \gamma_i}{n-2}\right\}\right\}. \tag{4.39}$$

Proof. By 4.1.4 Lemma 4.1.1 and (eq. (4.2)), we have the following LP problem.

$$\begin{aligned} f_i^{3,\mathrm{LP}}(x) = & \min_{a_i(k)\colon k=1,2,k_3} \frac{a_i(k_3)}{k_3} + \frac{x}{n} \\ & s.t.\ a_i(1) + a_i(2) + a_i(k_3) = \alpha_i - x, \\ & \quad a_i(1) + 2a_i(2) + k_3 a_i(k_3) = \beta_i - nx, \\ & \quad a_i(1) + 4a_i(2) + k_3^2 a_i(k_3) = \gamma_i - n^2 x, \\ & \quad a_i(1), a_i(2), a_i(k_3) \geqslant 0. \end{aligned} \tag{4.40}$$

Solving the linear system of 3 equations, we get

$$a_i(1) = \frac{2k_3(\alpha_i - x) - (k_3+2)(\beta_i - nx) + (\gamma_i - n^2x)}{k_3 - 1} \geqslant 0, \tag{4.41}$$

$$a_i(2) = \frac{-k_3(\alpha_i - x) + (k_3+1)(\beta_i - nx) - (\gamma_i - n^2x)}{k_3 - 2} \geqslant 0, \tag{4.42}$$

$$a_i(k_3) = \frac{2(\alpha_i - x) - 3(\beta_i - nx) + (\gamma_i - n^2x)}{(k_3 - 1)(k_3 - 2)} \geqslant 0. \tag{4.43}$$

Now the minimum of the objective function $f_i^{3,\mathrm{LP}}(x)$ becomes

$$f_i^{3,\mathrm{LP}}(x) = \frac{2(\alpha_i - x) - 3(\beta_i - nx) + (\gamma_i - n^2x)}{k_3(k_3 - 1)(k_3 - 2)} + \frac{x}{n}. \tag{4.44}$$

Since by assumption $a_i(k_3) \geqslant 0$, then $2(\alpha_i - x) - 3(\beta_i - nx) + (\gamma_i - n^2x) \geqslant 0$. Hence, $f_i^{3,\mathrm{LP}}(x)$ is decreasing in k_3 and we can minimise it by choosing $k_3 = n-1$. Hence, we get the following expression for $f_i^{3,\mathrm{LP}}(x)$.

$$f_i^{3,\mathrm{LP}}(x) = \frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)(n-3)} - \frac{3x}{n(n-3)} \tag{4.45}$$

The expressions for $a_i(1), a_i(2), a_i(k_3)$ simplifies to

$$a_i(1) = \frac{2(n-1)(\alpha_i - x) - (n+1)(\beta_i - nx) + (\gamma_i - n^2x)}{n-2} \geqslant 0, \tag{4.46}$$

$$a_i(2) = \frac{-(n-1)(\alpha_i - x) + n(\beta_i - nx) - (\gamma_i - n^2x)}{n-3} \geqslant 0, \tag{4.47}$$

$$a_i(n-1) = \frac{2(\alpha_i - x) - 3(\beta_i - nx) + (\gamma_i - n^2x)}{(n-2)(n-3)} \geqslant 0. \tag{4.48}$$

Note that $f_i^{3,\mathrm{LP}}(x)$ is decreasing in x so we need to choose the largest x possible. The 3 inequality constraints $a_i(1), a_i(2), a_i(n-1) \geqslant 0$ impose 3 classes of conditions on feasible values of x.

$$a_i(1) \geqslant 0 \implies x \leqslant \frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)} \tag{4.49}$$

$$a_i(2) \geqslant 0 \implies x \geqslant \frac{(n-1)\alpha_i - n\beta_i + \gamma_i}{n-1} \tag{4.50}$$

$$a_i(n-1) \geqslant 0 \implies x \leqslant \frac{2(n-1)\alpha_i - (n+1)\beta_i + \gamma_i}{n-2} \tag{4.51}$$

Since these constraints need to hold for $i = 1, ..., n$, combining them we get the following condition imposed on x.

$$x_{\min} \leqslant x \leqslant x_{\max} \tag{4.52}$$

where

$$x_{\min} = \max_{i=1,\dots,n} \left\{ \frac{(n-1)\alpha_i - n\beta_i + \gamma_i}{n-1} \right\}, \tag{4.53}$$

$$\begin{aligned} x_{\max} = \min \Bigg\{ & \min_{i=1,\dots,n} \left\{ \frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)} \right\}, \\ & \min_{i=1,\dots,n} \left\{ \frac{2(n-1)\alpha_i - (n+1)\beta_i + \gamma_i}{n-2} \right\} \Bigg\}. \end{aligned} \tag{4.54}$$

Let δ be the optimal choice of x that minimises $f_i^{3,\mathrm{LP}}(x)$. Then $\delta = x_{\max}$ since we need to maximise x. Substituting the result into $f_i^{3,\mathrm{LP}}(x)$, we get the following bound

$$\mathbb{P}(\mu \geqslant 3) \geqslant \ell_{3,\mathrm{LP}} = \sum_{i=1}^{n} \frac{2\alpha_i - 3\beta_i + \gamma_i}{(n-1)(n-2)(n-3)} - \frac{3\delta}{n(n-3)}. \tag{4.55}$$

$\square$

Remark 4.1.6. *Comparing the bound to the bound by Boros and Prékopa [10] using the aggregated partial information S_1, S_2, S_3, this bound $\ell_{3,LP}$ is sharper since it uses the disaggregated partial information equivalent to S_1^i, S_2^i, S_3^i, $i = 1, \dots, n$ with the same reasoning given in Remark 2.1.10.*

Bibliography

[1] András Prékopa. “Sharp bounds on probabilities using linear programming”. In: *Operations Research* 38.2 (1990), pp. 227–239.

[2] András Prékopa and Linchun Gao. “Bounding the probability of the union of events by aggregation and disaggregation in linear programs”. In: *Discrete applied mathematics* 145.3 (2005), pp. 444–454.

[3] Mauro Cerasoli. “On the probability of a Boolean polynomial of events”. In: *Discrete Mathematics* 44.2 (1983), pp. 221–227.

[4] Endre Boros and Joonhee Lee. “Boole’s probability bounding problem, linear programming aggregations, and nonnegative quadratic pseudo-Boolean functions”. In: *arXiv preprint arXiv:2110.10672* (2021).

[5] Jun Yang, Fady Alajaji, and Glen Takahara. “A short survey on bounding the union probability using partial information”. In: *arXiv preprint arXiv:1710.07576* (2017).

[6] Feng Qiu, Shabbir Ahmed, and Santanu S Dey. “Strengthened bounds for the probability of k-out-of-n events”. In: *Discrete Applied Mathematics* 198 (2016), pp. 232–240.

[7] D De Caen. “A lower bound on the probability of a union”. In: *Discrete mathematics* 169.1-3 (1997), pp. 217–220.

[8] Janos Galambos and Italo Simonelli. *Bonferroni-type inequalities with applications*. Springer, 1996.

[9] DA Dawson and D Sankoff. “An inequality for probabilities”. In: *Proceedings of the American Mathematical Society* 18.3 (1967), pp. 504–507.

[10] Endre Boros and András Prékopa. “Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur”. In: *Mathematics of Operations Research* 14.2 (1989), pp. 317–342.

[11] Hongyan Kuai, Fady Alajaji, and Glen Takahara. “A lower bound on the probability of a finite union of events”. In: *Discrete Mathematics* 215.1-3 (2000), pp. 147–158.

[12] Jun Yang, Fady Alajaji, and Glen Takahara. “Lower bounds on the probability of a finite union of events”. In: *SIAM Journal on Discrete Mathematics* 30.3 (2016), pp. 1437–1452.

[13] Munevver M Subasi et al. “New bounds for the probability that at least k-out-of-n events occur with unimodal distributions”. In: *Discrete Applied Mathematics* 226 (2017), pp. 138–157.

[14] Ersoy Subasi et al. “Sharp bounds for the probability that exactly k-out-of-n events occur with unimodal distributions”. In: *Annals of Operations Research* (2018), pp. 1–25.
[15] Dennis Wackerly, William Mendenhall, and Richard L Scheaffer. *Mathematical statistics with applications*. Cengage Learning, 2014.
[16] Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*. Vol. 6. Athena scientific Belmont, MA, 1997.
[17] L Richard Turner. *Inverse of the Vandermonde matrix with applications*. Tech. rep. 1966.
[18] David A. Cox, John Little, and Donal O'Shea. “Invariant Theory of Finite Groups”. In: *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer International Publishing, 2015, pp. 345–383.
[19] Michael Sullivan. *Algebra and Trigonometry*. Prentice Hall, 1996.