INS AND OUTS OF INCLUSION–EXCLUSION

R.E. Hartwig, Min Kang
North Carolina State University
e-mail: hartwig@math.ncsu.edu
kang@math.ncsu.edu

Abstract Inclusion-Exclusion identities and inequalities are obtained for valuations. Applications to cardinality, probability, max/min and least common multiples are presented.

keywords: Inclusion-Exclusion principle; Boole's inequality; Bonferroni inequality.

1 Introduction

In this treatise we shall examine the identities and inequalities associated with a valuation on a set $(S, +, \cdot)$. For finite sets these give the "inclusion-exclusion" (InEx) formulae and inequalities, while for probability the former yields the Poincaré's formula.

The basic valuation formula can directly be applied to such non-negative functions as content or measure (volume, area, length, weight, size, probability), as well as dimension, min/max and gcd/lcm. In the former case it is also possible to use indicator functions followed by the taking of a suitable linear functional, such as expected value, which will get us back to probability. When f is multiplicative, an easier and non-inductive way of proving the valuation formula is by using the symmetric functions of the roots $f(a_i)$ of a suitable polynomial.

Throughout the paper, $(S, +, \cdot)$ is a set S with two binary operations "$+$" and "$\cdot$", which are commutative and associative, and we further assume multiplicative idempontency $a \cdot a = a$ for all $a \in S$ and the distributive law $a \cdot (b + c) = a \cdot b + a \cdot c$, for all $a, b, c \in S$. The target space is a set $(T, \oplus, \otimes)$ with binary operations $\oplus$ and $\otimes$.

In general, we do not assume idempotency for addition, but when we do, we will clearly state that $a + a = a$ for all $a \in S$ as well. For instance, $(S, +, \cdot)$ can be a commutative distributive complemented lattice with $a \cdot b = a \wedge b = glb(a, b)$ and $a + b = a \vee b = lub(a, b)$. The most important of these is the power set $(\mathcal{P}(X), \cup, \cap)$.

Definition 1.1. *A function* $f: (S, +, \cdot) \to (T, \oplus, \otimes)$ *is called an* α-valuation *if*

$$f(a+b)\oplus[\alpha\otimes f(a\cdot b)] = f(a)\oplus f(b), \tag{1}$$

where $\alpha \in T$ *and* $a, b \in S$.

We shall examine the interplay between an α-valuation f and the operations of $+$ and $(\cdot)$ defined on S, and $\oplus$ and $\otimes$ defined on T. When there is no risk of ambiguity we shall as always write this as $f(a+b)+\alpha f(ab) = f(a)+f(b)$, in which the "multiplicative dot" has been dropped.

When T admits an additive inverse, we may rewrite this as $f(a+b) = f(a)+f(b)-\alpha f(ab)$. Needless to say when α is absent or $\alpha = 1$, we have a more symmetric unit-valuation. When this is the case, we do not need to define multiplication on T.

We shall primarily be interested in the case where (T,$\oplus$,$\otimes$) is $\mathbb{R}^+$, and $\alpha = 1+\varepsilon \geq 1$ (see [2]).

We shall examine additive as well as multiplicative results for α-valuations. Non-zero values of ε are used for example in the **exclusive-or** case (addition on a powerset), where $\alpha = 2$ and $f(a+a) = f(\emptyset) = 0$. It should be noted that $f(a+a) = (2-\alpha)f(a)$ and hence we have $f(a+a) = f(a)$ iff $\alpha = 1$.

2 Additive Results for α-valuations

We shall first need several definitions and notations dealing with triangular "slices".

For a given set $S \subseteq \mathbb{N}$, with $\#(S) = M$, we introduce the associated collection of lists.

Definition 2.1. *For* $k \leq M$,

$$V_k^S = \{(i_1, \ldots, i_k); i_1 < i_2 < \cdots < i_k, i_r \in S,\ \forall r = 1, \ldots, k\}.$$

We may alternatively think of this as the collection of all $\binom{M}{k}$ combinations of the M objects in S, taken k at a time. For example $V_2^{2,3,4} = \{(2,3),(2,4),(3,4)\}$. Initially we shall focus on $S = \{1, 2, \ldots, n\}$, and shorten $V_k^{1,2,\ldots,n}$ to $U_k^{(n)}$. In particular, $U_n^{(n)} = \{(1, 2, \ldots, n)\}$ is made of a single string.

Now let $A = \{x_1, \ldots, x_n\}$ be a collection of symbols. Then for $a \in A$ and $V \subseteq A^k$ we define $(V, a) = \{(x_1, \ldots, x_k, a); (x_1, \ldots, x_k) \in V\}$. We now have

Lemma 2.1. $U_k^{(n+1)} = U_k^{(n)} \cup (U_{k-1}^{(n)}, n+1)$

Proof. This is nothing but a partitioning of $U_k^{(n+1)}$ into terms that do or do not contain the highest index $n+1$. $\square$

Next, we introduce a collection of functions $g_k : A^k \to (T, \oplus)$, $k = 1, \dots, n$, where T is a suitable set with addition $\oplus$. For each $k = 1, 2, \dots, n$ we now define

Definition 2.2. $\bigoplus\limits_{U_k^{(n)}} g_k(x_{i_1}, \dots, x_{i_k}) = \bigoplus\limits_{1 \le i_1 < i_2 < \cdots < i_k \le n} g_k(x_{i_1}, x_{i_2}, \dots, x_{i_k})$, *where the addition is in T.*

We may now state:

Corollary 2.1.

(i) $$\bigoplus_{U_k^{(n+1)}} g_k(x_{i_1}, \dots, x_{i_k}) = \bigoplus_{U_k^{(n)}} g_k(x_{i_1}, \dots, x_{i_k}) \oplus \bigoplus_{U_{k-1}^{(n)}} g_k(x_{i_1}, \dots, x_{i_{k-1}}, n+1).$$

(ii) $$\bigoplus_{k=1}^{n+1} \bigoplus_{U_k^{(n+1)}} g_k(x_{i_1}, \dots, x_{i_k})$$

$$= \bigoplus_{k=1}^{n} \bigoplus_{U_k^{(n)}} g_k(x_{i_1}, \dots, x_{i_k}) \oplus \bigoplus_{k=1}^{n+1} \bigoplus_{U_{k-1}^{(n)}} g_k(x_{i_1}, \dots, x_{i_{k-1}}, n+1).$$

Note that in the second summation the term with $k = n+1$ is absent as $U_{n+1}^{(n)} = \emptyset$.

As a special case, we choose $g_k(a_1, \dots, a_k) = (-\alpha)^k f(a_1 \cdots a_k)$, where f is an α-valuation from $(S, +, \cdot)$ to $T = \mathbb{R}^+$ evaluated at the **product** of a_i. We further let $\mathbf{a} = (a_1, a_2, \dots)$ be a sequence of elements from S and for any $b \in S$ we set b$\mathbf{a}$= $(ba_1, ba_2, \dots)$.

For convenience we now also define the following "symmetric functions"

$$\sigma_k^{(n)}(\mathbf{a}, f) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} f(a_{i_1} a_{i_2} \dots a_{i_k}).$$

Note that $\sigma_n^{(n)}(\mathbf{a}, f) = f(a_1 \cdots a_n)$. When there is no risk of confusion we shall drop the brackets in the superscript and use σ_k^n instead of $\sigma_k^{(n)}$.

Again, if f is an α-valuation from $(S, +, \cdot)$ to $\mathbb{R}^+$ then we may recast Corollary (2.1) in terms of f as

Corollary 2.2.

$$(i)\quad \sigma_k^{n+1}(\boldsymbol{a},f)=\sigma_k^n(\boldsymbol{a},f)+\sigma_{k-1}^n(b\boldsymbol{a},f),$$

where $k=1,2,\dots,n+1$ *and* $b=a_{n+1}$.

$$(ii)\quad \sum_{k=1}^{n+1}(-\alpha)^{k-1}\sigma_k^{n+1}(\boldsymbol{a},f)=\sum_{k=1}^{n}(-\alpha)^{k-1}\sigma_k^n(\boldsymbol{a},f)+\sum_{k=1}^{n+1}(-\alpha)^{k-1}\sigma_{k-1}^n(b\boldsymbol{a},f)\,.$$

For consistency, we let $\sigma_0^n(b\mathbf{a},f)=f(b)$ for $k=1$ in part (i) of Corollary (2.1) and also note that $\sigma_k^n(\mathbf{a},f)$ is undefined for $k>n$.

We shall primarily be interested in the special sums,

$$S_n=S_n(\mathbf{a},f)=f(\sum_{i=1}^n a_i),$$

where $\mathbf{a}=(a_1,a_2,\dots)$. Using (1), we observe that

Lemma 2.2. *If* $b=a_{n+1}$, *then*

$$S_{n+1}(\boldsymbol{a})+\alpha S_n(b\boldsymbol{a})=S_n(\boldsymbol{a})+f(b) \tag{2}$$

Proof. $f(\sum_{i=1}^n a_i+b)+\alpha f[(\sum_{i=1}^n a_i)b]=f(\sum_{i=1}^n a_i)+f(b).$ □

Our aim is to solve the recurrence (2) for S_n. When T does not admit an additive inverse, we have to separate the even and odd values of n. We shall first introduce the following summations.

Definition 2.3. *(i)* $\lambda_k^n=\sum_{i=1}^k \alpha^{2i-2}\sigma_{2i-1}^n,\quad 2k-1\le n,$

(ii) $\mu_k^n=\sum_{i=1}^k \alpha^{2i-1}\sigma_{2i}^n,\quad 2k\le n.$

In this we dropped the superscript braces for convenience. For example,

$$\lambda_1^n=\sigma_1^n,\quad \lambda_2^n=\sigma_1^n+\alpha^2\sigma_3^n,\quad \lambda_3^n=\sigma_1^n+\alpha^2\sigma_3^n+\alpha^4\sigma_5^n$$

and

$$\mu_1^n=\alpha\sigma_2^n,\quad \mu_2^n=\alpha\sigma_2^n+\alpha^3\sigma_4^n,\quad \mu_3^n=\alpha\sigma_2^n+\alpha^3\sigma_4^n+\alpha^5\sigma_6^n.$$

Here μ_0^n is not defined, or it can be taken as the additive identity in T, if it exists. Before we can use induction we need the following identities.

Lemma 2.3. *Let* $b=a_{n+1}$, *then the following identities among* $\{\lambda_r^n\}$ *and* $\{\mu_r^n\}$ *hold.*

(i) $\lambda_r^{n+1}(\boldsymbol{a}) = \lambda_r^n(\boldsymbol{a}) + \alpha\mu_{r-1}^n(b\boldsymbol{a}) + f(b)$.

(ii) $\lambda_{r+1}^{n+1}(\boldsymbol{a}) = \lambda_r^n(\boldsymbol{a}) + \alpha\mu_r^n(b\boldsymbol{a}) + f(b) + \alpha^{2r}\sigma_{2r+1}^n(\boldsymbol{a})$.

(iii) $\mu_r^{n+1}(\boldsymbol{a}) = \mu_r^n(\boldsymbol{a}) + \alpha\lambda_r^n(b\boldsymbol{a})$.

(iv) $\mu_r^{n+1}(\boldsymbol{a}) + \alpha^{2r+1}\sigma_{2r+1}^n(b\boldsymbol{a}) = \mu_r^n(\boldsymbol{a}) + \alpha\lambda_{r+1}^n(b\boldsymbol{a})$.

Proof. Note that $b = a_{n+1}$ throughout.

(i) $\lambda_r^n(\mathbf{a})+\alpha\mu_{r-1}^n(b\mathbf{a})+f(b) = \sum_{k=1}^{r} \alpha^{2k-2}\sigma_{2k-1}^n(\mathbf{a})+\alpha\{\sum_{k=1}^{r-1} \alpha^{2k-1}\sigma_{2k}^n(b\mathbf{a})\}+ f(b) = [\sigma_1^n(\mathbf{a}) + f(b)] + \sum_{k=2}^{r} \alpha^{2k-2}\sigma_{2k-1}^n(\mathbf{a}) + \sum_{k=1}^{r-1} \alpha^{2k}\sigma_{2k}^n(b\mathbf{a}) = \sigma_1^{n+1}(\mathbf{a}) + \sum_{t=1}^{r-1} \alpha^{2t}\sigma_{2t+1}^n(\mathbf{a}) + \sum_{k=1}^{r-1} \alpha^{2k}\alpha_{2k}^n(b\mathbf{a}) = \sigma_1^{n+1}(\mathbf{a}) + \sum_{t=1}^{r-1} \alpha^{2t}[\sigma_{2t+1}^n(\mathbf{a}) + \sigma_{2t}^n(b\mathbf{a})] = \sum_{k=0}^{r-1} \alpha^{2k}\sigma_{2k+1}^{n+1}(\mathbf{a}) = \sum_{s=1}^{r} \alpha^{2s-2}\sigma_{2s-1}^{n+1}(\mathbf{a}) = \lambda_r^{n+1}(\mathbf{a})$.

(ii) $\lambda_r^n(\mathbf{a}) + \alpha\mu_r^n(b\mathbf{a}) + f(b) + \alpha^{2r}\sigma_{2r+1}^n(\mathbf{a}) = \lambda_r^n(\mathbf{a}) + \alpha\mu_{r-1}^n(b\mathbf{a}) + f(b) + \alpha^{2r}\sigma_{2r+1}^n(\mathbf{a})+\alpha^{2r}\sigma_{2r}^n(b\mathbf{a}) = \lambda_r^{n+1}(\mathbf{a})+\alpha^{2r}\sigma_{2r}^n(b\mathbf{a})+\alpha^{2r}\sigma_{2r+1}^n(\mathbf{a}) = \lambda_r^{n+1}(\mathbf{a})+ \alpha^{2r}\sigma_{2r+1}^{n+1}(\mathbf{a}) = \lambda_{r+1}^{n+1}(\mathbf{a})$, which proves (ii).

(iii) The right hand side equals $\sum_{k=1}^{r} \alpha^{2k-1}\sigma_{2k}^n(\mathbf{a})+\alpha \sum_{k=1}^{r} \alpha^{2k-2}\sigma_{2k-1}^n(b\mathbf{a}) = \sum_{k=1}^{r} \alpha^{2k-1}[\sigma_{2k}^n(\mathbf{a}) + \sigma_{2k-1}^n(b\mathbf{a})] = \sum_{k=1}^{r} \alpha^{2k-1}\sigma_{2k}^{n+1}(\mathbf{a}) = \mu_r^{n+1}(\mathbf{a})$.

(iv) From the right hand side, we have $\sum_{k=1}^{r} \alpha^{2k-1}\sigma_{2k}^n(\mathbf{a}) + \alpha\sum_{k=1}^{r+1} \alpha^{2k-2}\sigma_{2k-1}^n(b\mathbf{a}) = \sum_{k=1}^{r} \alpha^{2k-1}[\sigma_{2k}^n(\mathbf{a}) + \sigma_{2k-1}^n(b\mathbf{a})] + \alpha^{2r+1}\sigma_{2r+1}^n(b\mathbf{a})$, which is the left hand side. □

It should be noted that when $2r \geq n$, then σ_{2r+1}^n is absent. Moreover by part (iii) and (iv) in Lemma 2.3, we also have

Corollary 2.3. *(i)* $\mu_m^{2m+2}(\boldsymbol{a}) = \mu_m^{2m+1}(\boldsymbol{a}) + \alpha\lambda_m^{2m+1}(b\boldsymbol{a})$

(ii) $\mu_m^{2m+2}(\boldsymbol{a}) + \alpha^{2m+1}\sigma_{2m+1}^{2m+1}(b\boldsymbol{a}) = \mu_m^{2m+1}(\boldsymbol{a}) + \alpha\lambda_{m+1}^{2m+1}(b\boldsymbol{a})$.

We are now ready for the following identities on $S_n(\mathbf{a})$.

Theorem 2.1.

$$(i)\ S_{2m}(\boldsymbol{a}) + \mu_m^{2m}(\boldsymbol{a}) = \lambda_m^{2m}(\boldsymbol{a}) \quad \textit{for all } m \in \mathbb{N}\,, \tag{3}$$

$$(ii)\ S_{2m+1}(\boldsymbol{a}) + \mu_m^{2m+1}(\boldsymbol{a}) = \lambda_{m+1}^{2m+1}(\boldsymbol{a}) \quad \textit{for all } m \in \mathbb{N}\cup\{0\}\,. \tag{4}$$

Proof. (a) Both results are clearly true for initial values of m (in other words, $m = 1$ for part (i) and $m = 0$ for part (ii)). So let us assume that the result is true for $n = 2m$, and then show it for $n = 2m + 1$. Consider (2) with $n = 2m$ and add $\mu_m^{2m}(\mathbf{a})$ and $\alpha\mu_m^{2m}(b\mathbf{a})$ to both sides. This gives

$$S_{2m+1}(\mathbf{a})+\alpha[S_{2m}(b\mathbf{a})+\mu_m^{2m}(b\mathbf{a})]+\mu_m^{2m}(\mathbf{a}) = S_{2m}(\mathbf{a})+\mu_m^{2m}(\mathbf{a})+f(b)+\alpha\mu_m^{2m}(b\mathbf{a})$$

which by induction hypothesis reduces to

$$S_{2m+1}(\mathbf{a}) + \alpha\lambda_m^{2m}(b\mathbf{a}) + \mu_m^{2m}(\mathbf{a}) = \lambda_m^{2m}(\mathbf{a}) + f(b) + \alpha\mu_m^{2m}(b\mathbf{a}).$$

This in turn can be reduced by part (ii) and (iii) in Lemma (2.3) to

$$S_{2m+1}(\mathbf{a}) + \mu_m^{2m+1}(\mathbf{a}) = \lambda_{m+1}^{2m+1}(\mathbf{a}) - \alpha^{2m}\sigma_{2m+1}^{2m}(\mathbf{a}) = \lambda_{m+1}^{2m+1}(\mathbf{a}).$$

(b) Next we assume that the result holds for $n = 2m + 1$ and we will show it holds for $n = 2m + 2$. Consider (2) with $n = 2m + 1$ and $b = a_{2m+2}$, and add $\mu_m^{2m+1}(\mathbf{a})$ and $\alpha\mu_m^{2m+1}(b\mathbf{a})$ to both sides. This gives

$$\begin{aligned} & S_{2m+2}(\mathbf{a}) + \alpha[S_{2m+1}(b\mathbf{a}) + \mu_m^{2m+1}(b\mathbf{a})] + \mu_m^{2m+1}(\mathbf{a}) = \\ = \quad & S_{2m+1}(\mathbf{a}) + \mu_m^{2m+1}(\mathbf{a}) + f(b) + \alpha\mu_m^{2m+1}(b\mathbf{a}) \end{aligned}$$

which by induction hypothesis reduces to

$$S_{2m+2}(\mathbf{a}) + \alpha\lambda_{m+1}^{2m+1}(b\mathbf{a}) + \mu_m^{2m+1}(\mathbf{a}) = \lambda_{m+1}^{2m+1}(\mathbf{a}) + \alpha\mu_m^{2m+1}(b\mathbf{a}) + f(b)\,.$$

On account of Lemma (2.3) part (iv) and Corollary (2.3), we see that

$$S_{2m+2}(\mathbf{a}) + \mu_m^{2m+2}(\mathbf{a}) + \alpha^{2m+1}\sigma_{2m+1}^{2m+1}(b\mathbf{a}) = \lambda_{m+1}^{2m+2}(\mathbf{a}).$$

Lastly, note that $\sigma_{2m+1}^{2m+1}(b\mathbf{a}) = \sigma_{2m+2}^{2m+2}(\mathbf{a})$ with $b = a_{2m+2}$, and so we reach

$$S_{2m+2}(\mathbf{a}) + \mu_{m+1}^{2m+2}(\mathbf{a}) = \lambda_{m+1}^{2m+2}(\mathbf{a}).$$ □

If T admits additive inverses, we have the following valuation formulæ.

Corollary 2.4. *(i)* $S_n(\boldsymbol{a}) = f(\sum_{i=1}^{n} a_i) = \sum_{k=1}^{n}(-\alpha)^{k-1}\sigma_k^n(\boldsymbol{a}, f)$ (5)

$$(ii) \quad S_{2m}(\boldsymbol{a}) = \lambda_m^{2m}(\boldsymbol{a}) - \mu_m^{2m}(\boldsymbol{a})$$
$$(iii) \quad S_{2m+1}(\boldsymbol{a}) = \lambda_{m+1}^{2m+1}(\boldsymbol{a}) - \mu_m^{2m+1}(\boldsymbol{a})$$

Proof. We may list the first few sums in which $\sigma_k^n = \sigma_k^n(\mathbf{a})$:

$$\begin{array}{lclcl} S_2 & + & \alpha\sigma_2^2 & = & \sigma_1^2 \\ S_3 & + & (\alpha\sigma_2^3) & = & (\sigma_1^3 + \alpha^2\sigma_3^3) \\ S_4 & + & (\alpha\sigma_2^4 + \alpha^3\sigma_4^4) & = & (\sigma_1^4 + \alpha^2\sigma_3^4) \\ S_5 & + & (\alpha\sigma_2^5 + \alpha^3\sigma_4^5) & = & (\sigma_1^5 + \alpha^2\sigma_3^5 + \alpha^4\sigma_5^5) \end{array}$$

$$\begin{array}{lcl} S_6 & + & (\alpha\sigma_2^6 + \alpha^3\sigma_4^6 + \alpha^5\sigma_6^6) & = & (\sigma_1^6 + \alpha^2\sigma_3^6 + \alpha^4\sigma_5^6) \\ S_7 & + & (\alpha\sigma_2^7 + \alpha^3\sigma_4^7 + \alpha^5\sigma_6^7) & = & (\sigma_1^7 + \alpha^2\sigma_3^7 + \alpha^4\sigma_5^7 + \alpha^6\sigma_7^7)\,. \end{array}$$

More generally,

$$S_{2m} + (\alpha\sigma_2^{2m} + \alpha^3\sigma_4^{2m} + \cdots + \alpha^{2m-1}\sigma_{2m}^{2m}) = \sigma_1^{2m} + \alpha^2\sigma_3^{2m} + \cdots + \alpha^{2m-2}\sigma_{2m-1}^{2m}$$

and

$$\begin{aligned} S_{2m+1} + (\alpha\sigma_2^{2m+1} + \alpha^3\sigma_4^{2m+1} + \cdots + \alpha^{2m-1}\sigma_{2m}^{2m+1}) \\ = \sigma_1^{2m+1} + \alpha^2\sigma_3^{2m+1} + \cdots + \alpha^{2m}\sigma_{2m+1}^{2m+1}\,. \end{aligned} \qquad \square$$

Examples If an additive inverse exists, we may write

(i) $f(a+b+c) = f(a)+f(b)+f(c) - \alpha[f(ab)+f(ac)+f(bc)]+\alpha^2 f(abc).$ (6)

(ii)
$$\begin{aligned} f(a+a) &= (2-\alpha)f(a) \\ f(a+b+b) &= f(a) + (2-\alpha)f(b) - (2-\alpha)\alpha f(ab). \end{aligned} \qquad (7)$$

3 The valuation inequalities

Throughout let f be an α-valuation from $(S,+,\cdot)$ to $\mathbb{R}^+$. We begin by considering the matrices $M = (m_{ij})$ and $N = (n_{ij})$ with

$$m_{ij} = \lambda_i^{2m} - \mu_j^{2m}, \qquad \text{for} \quad 1 \le i,j \le m$$

$$n_{ij} = \lambda_i^{2m+1} - \mu_j^{2m+1}, \qquad \text{for} \quad 1 \le i \le m+1, 1 \le j \le m.$$

Then

$$m_{ij} \le m_{i+1,j}, \quad m_{i,j+1} \le m_{ij} \quad \textit{and} \quad n_{ij} \ge n_{i,j+1}, \quad n_{i+1,j} \ge n_{ij}.$$

In other words, the elements in the matrix M as well as N **increase** from right to left and top to bottom. We shall next use induction to prove the valuation inequalities. We first assume the validity for even n and prove it for odd n, and then reverse the parity.

Theorem 3.1. *(a) For $n = 2m$ with $m \in \mathbb{N}$ and any $k \ge 1$,*

$(a_1) S_{2m} \le \sigma_1^{2m}(\boldsymbol{a}) - \alpha\sigma_2^{2m}(\boldsymbol{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m}(\boldsymbol{a}) = \lambda_{k+1}^{2m} - \mu_k^{2m},$ *if* $k+1 \le m,$ (8)

$(a_2) S_{2m} \ge \sigma_1^{2m}(\boldsymbol{a}) - \alpha\sigma_2^{2m}(\boldsymbol{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m}(\boldsymbol{a}) = \lambda_k^{2m} - \mu_k^{2m},$ *if* $k \le m.$ (9)

(b) For $n = 2m+1$ with $m \in \mathbb{N} \cup \{0\}$ and any $k \geq 1$,

$(b_1) S_{2m+1} \leq \sigma_1^{2m+1}(\boldsymbol{a}) - \alpha\sigma_2^{2m+1}(\boldsymbol{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m+1}(\boldsymbol{a}) = \lambda_{k+1}^{2m+1} - \mu_k^{2m+1}$, *if $k \leq m$,*

$(b_2) S_{2m+1} \geq \sigma_1^{2m+1}(\boldsymbol{a}) - \alpha\sigma_2^{2m+1}(\boldsymbol{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m+1}(\boldsymbol{a}) = \lambda_k^{2m+1} - \mu_k^{2m+1}$, *if $k \leq m$.*

Proof. (a) The inequalities clearly hold for the initial values of m (in other words, $m = 1$ for part (a) and $m = 0$ for part (b)), so let us assume they both hold for all values of $r \leq n$ and that $b = a_{n+1}$. Recall that

$$S_{n+1} = f(\sum_{i=1}^{n} a_i + b) = f(\sum_{i=1}^{n} a_i) + f(b) - \alpha f[\sum_{i=1}^{n}(a_i b)].$$

By induction hypothesis, in the first summation on the right hand side, we may apply the inequality from (8) with $2k+1$ terms and sequence $\mathbf{a}$, while in the second summation we apply (9) with 2k terms and sequence b$\mathbf{a}$, respectively. This gives

$S_{2m+1} \leq [\sigma_1^{2m}(\mathbf{a}) - \alpha\sigma_2^{2m}(\mathbf{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m}(\mathbf{a})] + f(b) - \alpha[\sigma_1^{2m}(b\mathbf{a}) - \alpha\sigma_2^{2m}(b\mathbf{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m}(b\mathbf{a})]$

$= [f(b) + \sigma_1^{2m}(\mathbf{a})] - \alpha[\sigma_2^{2m}(\mathbf{a}) + \sigma_1^{2m}(b\mathbf{a})] + \cdots + \alpha^{2k}[\sigma_{2k+1}^{2m}(\mathbf{a}) + \sigma_{2k}^{2m}(b\mathbf{a})]$

$= \sigma_1^{2m+1}(\mathbf{a}) - \alpha\sigma_2^{2m+1}(\mathbf{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m+1}(\mathbf{a})$.

On the other hand, if we apply (9) in the first summation on the right hand side, with $2k$ terms and sequence $\mathbf{a}$, and apply (8) with $2k-1$ terms and sequence $b\mathbf{a}$, then we obtain

$S_{2m+1} \geq [\sigma_1^{2m}(\mathbf{a}) - \alpha\sigma_2^{2m}(\mathbf{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m}(\mathbf{a})] + f(b) - \alpha[\sigma_1^{2m}(b\mathbf{a}) - \alpha\sigma_2^{2m}(b\mathbf{a}) + \cdots + \alpha^{2k-2}\sigma_{2k-1}^{2m}(b\mathbf{a})]$

$= [f(b) + \sigma_1^{2m}(\mathbf{a})] - \alpha[\sigma_2^{2m}(\mathbf{a}) + \sigma_1^{2m}(b\mathbf{a})] + \cdots - \alpha^{2k-1}[\sigma_{2k}^{2m}(\mathbf{a}) + \sigma_{2k-1}^{2m}(b\mathbf{a})]$

$= \sigma_1^{2m+1}(\mathbf{a}) - \alpha\sigma_2^{2m+1}(\mathbf{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m+1}(\mathbf{a})$.

Next, we assume that (b1) and (b2) hold for $r \leq 2m+1$ and use (a1) and (a2) in the α-evaluation formula for $2m+2$,

$$S_{2m+2}(\mathbf{a}) = S_{2m+1}(\mathbf{a}) + f(b) - \alpha S_{2m+1}(b\mathbf{a})$$

to give

$S_{2m+2} \leq [\sigma_1^{2m+1}(\mathbf{a}) - \alpha\sigma_2^{2m+1}(\mathbf{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m+1}(\mathbf{a})] + f(b) - \alpha[\sigma_1^{2m+1}(b\mathbf{a}) - \alpha\sigma_2^{2m+1}(b\mathbf{a}) + \cdots - \alpha^{2k-1}\sigma_{2k}^{2m+1}(b\mathbf{a})] = [f(b) + \sigma_1^{2m+1}(\mathbf{a})] - \alpha[\sigma_2^{2m+1}(\mathbf{a}) + \sigma_1^{2m+1}(b\mathbf{a})] + \cdots + \alpha^{2k}[\sigma_{2k+1}^{2m+1}(\mathbf{a}) + \sigma_{2k}^{2m+1}(b\mathbf{a})] = \sigma_1^{2m+2}(\mathbf{a}) - \alpha\sigma_2^{2m+2}(\mathbf{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m+2}(\mathbf{a})$.

Likewise, noting that $k+1 \le m+1$, we reach

$S_{2m+2} \ \ge \ [\sigma_1^{2m+1}(\mathbf{a}) - \alpha\sigma_2^{2m+1}(\mathbf{a}) + \cdots - \alpha^{2k+1}\sigma_{2k+2}^{2m+1}(\mathbf{a})] + f(b) - \alpha[\sigma_1^{2m+1}(b\mathbf{a}) - \alpha\sigma_2^{2m+1}(b\mathbf{a}) + \cdots + \alpha^{2k}\sigma_{2k+1}^{2m+1}(b\mathbf{a})] = [f(b) + \sigma_1^{2m+1}(\mathbf{a})] - \alpha[\sigma_2^{2m+1}(\mathbf{a}) + \sigma_1^{2m+1}(b\mathbf{a})] + \cdots - \alpha^{2k+1}[\sigma_{2k+2}^{2m+1}(\mathbf{a}) + \sigma_{2k+1}^{2m+1}(b\mathbf{a})] = \sigma_1^{2m+2}(\mathbf{a}) - \alpha\sigma_2^{2m+2}(\mathbf{a}) + \cdots - \alpha^{2k+1}\sigma_{2k+2}^{2m+2}(\mathbf{a})$,

which completes the proof. □

The InEx inequalities suggest that there should be inequalities relating the symmetric functions. We shall now show this, for the case where $\alpha f(ab) \le f(a)$ and $k+1 \le n \le 2k+1$. We first need some basic facts about the binomial sets. We shall denote the existence of an injective map from $V_{k+1}^{1,\ldots,n}$ into $V_k^{1,..n}$ by $V_{k+1}^{1,\ldots,n} \hookrightarrow V_k^{1,\ldots,n}$.

Lemma 3.1. *If $k+1 \le n \le 2k+1$ then we can find an injection (one-to-one map) from $V_{k+1}^{1,\ldots,n}$ into $V_k^{1,..n}$.*

Proof. The proof follows by induction on n, and is very similar to that of Theorem (3.1) in that we have to separate even and odd values of n. When $n=3$, the only possible values for k are $k=1$ or 2. The result is now easily seen for these cases because $\binom{3}{2} = \binom{3}{1}$ and the injection is supplied by taking the "complement" in $\{1,2,3\}$. On the other hand, because $\binom{3}{3} = 1$, we can drop any one of the 3 digits in $V_3^{1,2,3}$=$\{(1,2,3)\}$, to obtain a unique image in $V_2^{1,2,3}$. Let us now assume that $V_{k+1}^{1,2,\ldots,2m} \hookrightarrow V_k^{1,2,\ldots,2m}$ for all k such that $k+1 \le 2m \le 2k+1$, in other words, $k = m, m+1, \ldots, 2m-1$. Now we wish to show that $V_{k+1}^{1,2,\ldots,2m+1} \hookrightarrow V_k^{1,2,\ldots,2m+1}$ again for all k such that $k+1 \le 2m+1 \le 2k+1$, in other words, $k = m, \ldots, 2m$. We observe that $V_k^{1,2,\ldots,2m+1}$ can be written as a disjoint union, by separating the terms that start with a digit 1 from those that start with a digit 2 etc. Indeed, through this natural decomposition, we have a bijection between $V_{m+1}^{1,\ldots,2m+1}$ and $\hat{V}_m^{2,\ldots,2m+1} \cup \hat{V}_m^{3,\ldots,2m+1} \cup \cdots \cup \hat{V}_m^{m+2,\ldots,2m+1}$ where $\hat{V}_m^{a+1,\ldots,b} = \{(a, x_1, \ldots, x_m) \,|\, a+1 \le x_1 < \cdots < x_m \le b\} \subset V_{m+1}^{1,\ldots,b}$. Obviously there is a natural bijection between $\hat{V}_m^{a+1,\ldots,b}$ and $V_m^{a+1,\ldots,b}$ by dropping the first coordinate of the vectors. Hence we note that the number of list in $\hat{V}_m^{2,\ldots,2m+1}$ equals $\binom{2m}{m}$, while the number in $\hat{V}_m^{3,\ldots,2m+1}$ is $\binom{2m-1}{m}$ etc. By the hypothesis we can find injections from each of the $V_m^{r,r+1,\ldots,2m+1}$ into $V_{m-1}^{r,r+1,\ldots,2m+1}$ for all $r = 2, \ldots, m+2$. That is, we have, by the induction hypothesis,

$$V_m^{2,3,\ldots,2m+1} \hookrightarrow V_{m-1}^{2,3,\ldots,2m+1},$$

$$V_m^{3,4,\ldots,2m+1} \hookrightarrow V_{m-1}^{3,4,\ldots,2m+1},$$

$$\vdots$$

$$V_m^{m+2,\dots,2m+1} \hookrightarrow V_{m-1}^{m+2,\dots,2m+1}.$$

Combining these and the natural bijections between $\hat{V}_n^{a+1,\dots,b}$ and $V_n^{a+1,\dots,b}$, we have an injection from $V_{m+1}^{1,2,\dots,2m+1}$ into $\hat{V}_{m-1}^{2,3,\dots,2m+1} \cup \hat{V}_{m-1}^{3,4,\dots,2m+1} \cup \dots \cup \hat{V}_{m-1}^{m+2,\dots,2m+1}$ which is a subset of $V_m^{1,2,\dots,2m+1}$.

It similarly follows that $V_{m+2}^{1,2,\dots,2m+1} \hookrightarrow V_{m+1}^{1,2,\dots,2m+1}$ etc. Combining these and the natural bijections between $\hat{V}_n^{a+1,\dots,b}$ and $V_n^{a+1,\dots,b}$, the end result is that $V_{k+1}^{1,2,\dots,2m+1} \hookrightarrow V_k^{1,2,\dots,2m+1}$ for $k = m, \dots, 2m$. Hence this shows that if the result holds for even n then it also holds for the next odd n.

Let us now turn to the converse and assume that

$$V_{k+1}^{1,2,\dots,2m-1} \hookrightarrow V_k^{1,2,\dots,2m-1} \text{ for } k \geq m-1.$$

We wish to show that

$$V_{m+1}^{1,2,\dots,2m} \hookrightarrow V_m^{1,2,\dots,2m}, \quad V_{m+2}^{1,2,\dots,2m} \hookrightarrow V_{m+1}^{1,2,\dots,2m} \quad ,\dots, V_{2m}^{1,2,\dots,2m} \hookrightarrow V_{2m-1}^{1,2,\dots,2m}.$$

We focus on the first term with $k = m$, and again write it as a disjoint union of subsets, as

$$V_{m+1}^{1,2,\dots,2m} = \hat{V}_m^{2,\dots,2m} \cup \hat{V}_m^{3,\dots,2m} \cup \dots \cup \hat{V}_m^{m+1,\dots,2m}.$$

It again follows by the induction hypothesis that
$V_m^{2,\dots,2m} \hookrightarrow V_{m-1}^{2,\dots,2m}$, $V_m^{3,\dots,2m} \hookrightarrow V_{m-1}^{3,\dots,2m}$,$\dots$, and $V_m^{m+1,\dots,2m} \hookrightarrow V_{m-1}^{m+1,\dots,2m}$.
The end result is that $V_{m+1}^{1,2,\dots,2m} \hookrightarrow V_m^{1,2,\dots,2m}$. It similarly follows for the other pieces, giving the desired injection

$$V_{k+1}^{1,2,\dots,2m} \hookrightarrow V_k^{1,2,\dots,2m}$$

for all $k = m, \dots, 2m-1$, i.e. for all k such that $k+1 \leq 2m \leq 2k+1$. □

We can now capitalize on the existence of the injection to derive the following inequalities for the "symmetric" functions. In the following Corollary, we do not need to assume the existence of the additive inverse in T.

Corollary 3.1. *Suppose that $\alpha f(ab) \leq f(a)$. Then for $k+1 \leq n \leq 2k+1$,*

$$\sigma_k^n(\boldsymbol{a}) \geq \alpha \sigma_{k+1}^n(\boldsymbol{a}) \tag{10}$$

Proof. We first observe that $\#(\sigma_k^n(\mathbf{a})) = \binom{n}{k}$. Now because of the injection, we can associate to each term $f(a_{i_1}a_{i_2}\dots a_{i_{k+1}})$ in $\sigma^n_{k+1}(\mathbf{a})$ a distinct term $f(a_{j_1}\cdots a_{j_k})$ in $\sigma^n_k(\mathbf{a})$ obtained from the former by deleting **exactly one** of the a_{i_r}. Since $\alpha f(ab)\le f(a)$, we see that $\alpha f(a_{i_1}a_{i_2}\cdots a_{i_{k+1}})\le f(a_{j_1}\cdots a_{j_k})$ and hence that $\alpha\sigma^n_{k+1}(\mathbf{a})\le\sigma^n_k(\mathbf{a})$. $\square$

Remark For the case where $\alpha f(ab)\le f(a)$, half of the valuation inequalities follow from the local fact that $\sigma^n_k(\mathbf{a})\ge\alpha\sigma^n_{k+1}(\mathbf{a})$.

4 Additive valuation formula for multiplicative valuations with $f(ab)=f(a)f(b)$.

(a) When f is "multiplicative" i.e. $f(ab)=f(a)f(b)$, then the valuation formula can be simplified. In this case the symmetric functions simplify to

$$\sigma_k=\sigma_k(\mathbf{a},f)=\sum_{1\le i_1<i_2<\cdots<i_k\le n} f(a_{i_1})f(a_{i_2})\cdots f(a_{i_k}).$$

Then

$$\begin{aligned} f(\sum_{i=1}^n a_i) &= \sigma_1-\alpha\sigma_2+\alpha^2\sigma_3-\dots(-1)^{n-1}\alpha^{n-1}\sigma_n \\ &= \tfrac{1}{\alpha}[1-\prod_{i=1}^n(1-\alpha f(a_i))]. \end{aligned} \tag{11}$$

A particularly important example is that of the **indicator function** $\chi_A(s)$ of a set A. We shall examine these functions shortly in detail.

(b) On the other hand, if we have for each a in S a "complement" a' in S such that
(i) $(a+b)'=a'b'$ and (ii) $f(a')=1-\alpha f(a)$, then we can obtain (11) *directly* via

$$\begin{aligned} f(\sum_{i=1}^n a_i) &= \tfrac{1}{\alpha}[1-f[(\sum_{i=1}^n a_i)']]=\tfrac{1}{\alpha}[1-f[\prod_{i=1}^n a_i']] \\ &= \tfrac{1}{\alpha}[1-\prod_{i=1}^n f(a_i')]=\tfrac{1}{\alpha}[1-\prod_{i=1}^n[1-\alpha f(a_i)]] \end{aligned} \tag{12}$$

(c) For the case when $\alpha=1$, it suffices to have a "complement" a' such that

$$f(a)=f(ab')+f(ab).$$

4.1 Examples of Valuation equalities

Let us now turn to the various applications of the valuation formula.

Indeed, most of these deal with a collection of subsets of set X as the poset $(\mathcal{P}(X),\subseteq)$ and the lattice $(S,+,\cdot)=(\mathcal{P}(X),\cup,\cap)$, where in addition, f is some type of "content/size" such as volume, area, length or cardinality. We shall denote the collection of all **finite** subsets of set X, by $F(X)$.

Example 4.1. *Let* $(S,+,\cdot) = (F(X),\cup,\cap)$ *with* $f(A) = \#(A)$*, the cardinality of* A*. Then*

$$\#(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\le i_1<i_2<\cdots<i_k\le n} \#(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})\,.$$

Example 4.2. *Let* $I(\mathbb{R})$ *be an algebra or* σ*-algebra of intervals, with Lebesgue measure as length. Suppose* S *is a the collection of unions of intervals of* $\mathbb{R}$*, and let* $f(A)$ *be the length of* $\ell(A)$ *of* A*, so that* $(S,+,\cdot) = (I(\mathbb{R}),\cup,\cap)$*. Then*

$$\ell(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\le i_1<i_2<\cdots<i_k\le n} \ell(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})\,.$$

Example 4.3. $(S,+,\cdot) = (\mathcal{F},\cup,\cap)$*, where* $\mathcal{F} \subset \mathcal{P}(\Omega)$ *is a* σ*-field on* Ω *and* $(\Omega,\mathcal{F},\mathbb{P})$ *is a probability triple. We set* $f(A) = P(A)$*, the probability of* A*. Then* P *is a valuation and hence we have the Poincare formula:*

$$P(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\le i_1<i_2<\cdots<i_k\le n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})\,.$$

Needless to say, no indicator functions, expected values or integrals were needed to derive this.

Example 4.4. *Let* $(S,+,\cdot) = (\mathbb{R}^+,\max,\min)$ *and let* f *be the identity map. (This is not non-negative, so we have at first to restrict* S *to* $\mathbb{R}^+$*.)*

Now for two real numbers a *and* b*, it is easily seen that* $\max(a,b) = a + b - \min(a,b)$ *and* $\min\{a,\max(b,c)\} = \max\{\min(a,b),\min(a,c)\}$*, so that the identity map is a valuation. This gives*

$$\begin{aligned}\max\{a_1,\ldots,a_n\} \quad &= \quad (a_1+\cdots+a_n) - \sum_{i<j}\min\{a_i,a_j\} + \\ &+ \quad \sum_{i<j<k}\min\{a_i,a_j,a_k\} - \cdots + (-1)^{n-1}\min\{a_1,a_2,\ldots,a_n\}.\end{aligned}$$

Example 4.5. *Let* S *be a collection of all the subsets of* X *with* $(S,+,\cdot) = (\mathcal{P}(X),\cup,\cap)$*. For any prime* p*, and any valuation* f *with* $\alpha = 1$*, we have the multiplicative formula,*

$$p^{f(\cup_{i=1}^{n}A_i)}=\frac{\prod_{i=1}^{n}p^{f(A_i)}\cdot\prod_{1\le i<j<k\le n}p^{f(A_i\cap A_j\cap A_k)}\cdots\prod_{1\le i_1<\cdots<i_m\le n}p^{f(A_{i_1}\cap\cdots\cap A_{i_m})}\cdot}{\prod_{i<j}p^{f(A_i\cap A_j)}\cdots\prod_{1\le i_1<\cdots<i_\ell\le n}p^{f(A_{i_1}\cap\cdots\cap A_{i_\ell})}},$$

where $m:=\max\{k\le n\,|\,k \textit{ is odd}\}$ *and* $\ell:=\max\{k\le n\,|\,k \textit{ is even}\}$.

Example 4.6. $S=\mathbb{N}$ *and* $a\le b$ *if and only if* $a\mid b$.
Any positive integer can be expressed as a unique product of prime powers $a=p_1^{k_1}p_2^{k_2}\ldots p_r^{k_r}$, *with* $p_1<p_2<\ldots$. *If* b *is likewise expanded as* $b=p_1^{t_1}p_2^{t_2}\ldots p_r^{t_r}$, *then clearly* $a\mid b$ *iff* $k_i\le t_i$ *for all* i. *Moreover* $(a,b)=gcd(a,b)=\prod p_i^{min\{k_i,t_i\}}$ *and* $[a,b]=lcm(a,b)=\prod p_i^{max\{k_i,t_i\}}$. *Since min/max obey the InEx law on* $\mathbb{N}$, *we may conclude that the gcd and lcm satisfy the* multiplicative version *of the InEx law. In other words,*

$$[a_1,\ldots,a_n]=\frac{\prod_{i=1}^{n}a_i\cdot\prod_{1\le i_1<i_2<i_3\le n}(a_{i_1},a_{i_2},a_{i_3})\cdots\prod_{1\le i_1<\cdots<i_m\le n}(a_{i_1},\ldots,a_{i_m})}{\prod_{1\le i_1<i_2\le n}(a_{i_1},a_{i_2})\cdots\prod_{1\le i_1<\cdots<i_\ell\le n}(a_{i_1},\ldots,a_{i_\ell})},\tag{13}$$

where $m:=\max\{k\le n\,|\,k \textit{ is odd}\}$ *and* $\ell:=\max\{k\le n\,|\,k \textit{ is even}\}$. *For example,*

$$\begin{aligned}[a,b,c]&=\tfrac{abc(a,b,c)}{(a,b)(a,c)(b,c)}\quad and\\ [a,b,c,d]&=\tfrac{abcd(a,b,c)(a,b,d)(a,c,d)(b,c,d)}{(a,b)(a,c)(a,d)(b,c)(b,d)(c,d)(a,b,c,d)}\end{aligned}\tag{14}$$

This is a "multiplicative version" of the Inclusion Exclusion rule for an evaluation map f with $\alpha=1$ such as

$$f(a+b+c)=f(a)+f(b)+f(c)-[f(ab)+f(ac)+f(bc)]+f(abc).$$

Example 4.7. *Let* S *be the collection of all subspaces of a vector space* U, *with* $V+W$ *being the vector space direct sum and* $V\cdot W=V\cap W$. *Then* $f(V)=\dim(V)$ *is a valuation on* S *with* $\alpha=1$, *and*

$$dim(\sum_{k=1}^{n}V_k)=\sum_{k=1}^{n}(-1)^{k-1}\cdot\sum_{1\le i_1<i_2<\cdots<i_k\le n}\dim(V_{i_1}\cap V_{i_2}\cap\cdots\cap V_{i_k})$$

Example 4.8. $S=\mathbb{Z}_2^n=\{\boldsymbol{x}=(x_1,\ldots,x_n)\,|\,x_i\in\mathbb{Z}_2,\forall i=1,\ldots,n\}$ *with coordinate-wise addition and scalar multiplication. These operations are commutative, multiplication is idempotent and distributes over addition. The Hamming metric* $h(\boldsymbol{a})$ *counts the number of ones in the vector* $\boldsymbol{a}$. *It is a valuation that satisfies the "cosine rule"*

$$h(\boldsymbol{a}+\boldsymbol{b})=h(\boldsymbol{a})+h(\boldsymbol{b})-2h(\boldsymbol{a}.\boldsymbol{b}),\tag{15}$$

which is the InEx equation with $\alpha=2$.

Example 4.9. *$S = \mathbb{B}^n = \{\boldsymbol{x}; (x_1, \ldots, x_n) \,|\, x_i \in \mathbb{B}\,, \forall i = 1, \ldots, n\}$ with the Boolean scalar operations $x + y = x \vee y$ and $x \cdot y = x \wedge y$. The vector operations are again defined coordinate-wise. The Hamming metric is again a valuation, but this satisfies In-Ex equation with $\alpha = 1$, i.e.*

$$h(\boldsymbol{a} + \boldsymbol{b}) = h(\boldsymbol{a}) + h(\boldsymbol{b}) - h(\boldsymbol{a}.\boldsymbol{b}) \tag{16}$$

5 Examples of Valuation inequalities

Example 5.1. *$(S, +, \cdot) = (\mathcal{F}, \cup, \cap)$, where $\mathcal{F} \subset \mathcal{P}(\Omega)$ is a σ-field on Ω and $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability triple such that $f(A) = P(A)$, the probability of A. We have*

(i) $$P(\bigcup_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} P(A_i)\,, \quad \textit{(Boole's law)}$$

(ii) $$P(\bigcup_{k=1}^{n} A_k) \geq \sum_{k=1}^{n} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2})\,.$$

Example 5.2. *Next let $(S, +, \cdot) = (\mathbb{R}, \max, \min)$ and let f be the identity map (extended from $\mathbb{R}^+$). We have*

$\max\{a_1, \ldots, a_n\} \leq (a_1 + \cdots + a_n)\,,$
$\max\{a_1, \ldots, a_n\} \geq (a_1 + \cdots + a_n) - \sum_{i<j} \min\{a_i, a_j\}\,,$
$\max\{a_1, \ldots, a_n\} \leq (a_1 + \cdots + a_n) - \sum_{i<j} \min\{a_i, a_j\} + \sum_{i<j<k} \min\{a_i, a_j, a_k\}\,.$

Example 5.3. *Applying the above to prime powers we arrive at*

$[a, b, c, d] \leq abcd\,,$
$[a, b, c, d] \geq \frac{abcd}{(a,b)(a,c)(a,d)(b,c)(b,d)(c,d)}\,,$
$[a, b, c, d] \leq \frac{abcd}{(a,b)(a,c)(a,d)(b,c)(b,d)(c,d)}(a, b, c)(a, b, d)(a, c, d)(b, c, d),$

where we actually attain equality in the last inequality above.

6 Indicator functions

Many of the applications of the "inclusion-exclusion" formula using sets can be derived using the indicator functions (also called characteristic functions). The main purpose of introducing these functions is to convert manipulations involving sets into manipulations involving functions. i.e. turn a Boolean algebra into a Boolean ring. Let us now recap some of its properties.

Given a set S and a subset $A \subset S$, the indicator function of A, $\chi_A : S \to \{0,1\}$ is defined by

$$\chi_A(s) = \begin{cases} 1 & if s \in A \\ 0 & if\ s \notin A \end{cases}.$$

Closely related is the function

$$\phi_A(s) = 1 - 2\chi_A(s) = \begin{cases} 1 & if s \in A \\ -1 & if\ s \notin A \end{cases}.$$

Again, for subsets $A, B, C, \dots$ of S and $\cup, \cap$ and $\oplus$ stand for union, intersection and the operation, "exclusive or (XOR)", i.e. $A \oplus B = (A \cap B^c) \cup (B \cap A^c)$ and $A \oplus A = (A \cap A^c) \cup (A \cap A^c) = \emptyset$.

For real valued functions on S we define $f \leq g$ iff $f(s) \leq g(s)$ for all $s \in S$. We shall also abbreviate χ_{A_i} to χ_i, when necessary. Some of the valuation maps associated with indicator functions are demonstrated below.

1. $\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B} \quad \phi_{A\cup B} = \phi_A + \phi_B - \phi_{A\cap B}$ (1-valuations)

2. $\chi_{A\oplus B} = \chi_A + \chi_B - 2\chi_A\chi_B$ (2-valuation), $\quad \phi_{(A\oplus B)} = \phi_A.\phi_B$

3. χ provides a commutative valuation on $(\mathcal{P}(X), \cup, \cap)$ with $\alpha = 1$ and as such the derivation of (12) holds. It is traditional to derive this as

$$\chi_{\cup A_i} = 1 - \chi_{\cap A_i^c} = 1 - \prod_{i=1}^{n}(1 - \chi_{A_i}) = \sigma_1 - \sigma_2 + \dots + (-1)^{n-1}\sigma_n,$$

which rewrites as

$$\chi_{\cup A_i} = \sum_{1}^{n} \chi_i - \sum_{i<j} \chi_i\chi_j + \sum_{i<j<k} \chi_i\chi_j\chi_k + \dots + (-1)^{n-1}\chi_1\chi_2\dots\chi_n\,. \tag{17}$$

4. Since χ is a 2-valuation with $\alpha = 2$ on $(\mathcal{P}(X), \oplus, \cap)$ with $\oplus =$ XOR, we have from (11)

$$\chi_{\oplus_{i=1}^{n} A_i} = \sigma_1 - 2\sigma_2 + 4\sigma_3 + \dots + (-1)^{n-1}2^{n-1}\sigma_n.$$

5. $\sum\limits_{k=1}^{n}\prod\limits_{i\neq k}(1-\chi_i) = \sum\limits_{k=1}^{n}\prod\limits_{i\neq k}\chi_{A_i^c} = \sum\limits_{k=1}^{n}\chi_{(\bigcap\limits_{i\neq k} A_i^c)} = \sum\limits_{k=1}^{n}\chi_{(\bigcup\limits_{i\neq k} A_i)^c} = n - \sum\limits_{k=1}^{n}\chi_{\bigcup\limits_{i\neq k} A_i}$,
which parallels Lagrange interpolation.

6. Since indicator functions are multiplicative and act on sets, the valuation inequalities can also be derived using combinatorics. This is instructive in its own right. The valuation inequalities for indicator functions take the following forms,

   (i) $\chi_{\cup A_i} \leq \sum\limits_1^n \chi_i$

   (ii) $\chi_{\cup A_i} \geq \sum\limits_1^n \chi_i - \sum\limits_{i<j} \chi_i \chi_j$

   (iii) $\chi_{\cup A_i} \leq \sum\limits_1^n \chi_i - \sum\limits_{i<j} \chi_i \chi_j + \sum\limits_{i<j<k} \chi_i \chi_j \chi_k$

   (iv) $\chi_{\cup A_i} \geq \sum_{k=1}^r (-1)^{k-1} \cdot \sum\limits_{1 \leq i_1 < \cdots < i_k \leq n} \chi_{i_1} \chi_{i_2} \ldots \chi_{i_k}$ when r is even

   (v) $\chi_{\cup A_i} \leq \sum\limits_{k=1}^r (-1)^{k-1} \cdot \sum\limits_{1 \leq i_1 < \cdots < i_k \leq n} \chi_{i_1} \chi_{i_2} \ldots \chi_{i_k}$ when r is odd.

   To prove these using sets, it suffices to show the inequality at any point s in the union. Now the union $\cup A_i$ is made up of 2^n disjoint subsets (called atoms or petals) of the form $B_i = A_1 \cap A_2 \cdots \cap A_i \cap A_{i+1}^c \cdots \cap A_n^c$ etc. and it suffices to show that the inequalities hold for for some fixed s $\in B_i$. For the remaining subsets the result follows by symmetry. As an example we consider the case where $r = 3$.

   If $s \in B_1 = A_1 \cap A_2 \cdots \cap A_i \cap A_{i+1}^c \cdots \cap A_n^c$ then $s \in A_1 \cap A_2 \cdots \cap A_i$ and the left hand side which equals $\chi_{\cup A_i}(s)$ yields the value 1. On the other hand the right hand side gives the value: $\binom{i}{1} - \binom{i}{2} + \binom{i}{3}$. The result now follows from the binomial identity

   $$\sum_{k=0}^r (-1)^k \binom{i}{k} = \begin{cases} 0 & if\ r = i \\ \binom{i-1}{r}(-1)^r & if\ r < i. \end{cases} \tag{18}$$

   This yields the desired inequalities:
   (i) $\binom{n}{0} \geq \binom{n}{1} - \binom{n}{2} + \cdots + (-1)^{r-1}\binom{n}{r}$ when r is even and
   (ii) $\binom{n}{0} \leq \binom{n}{1} - \binom{n}{2} + \ldots + (-1)^{r-1}\binom{n}{r}$ when r is odd.

7 Multiplicative Formulae for α-valuations

Let f be an α-valuation on $(S, +, \cdot)$. We begin by noting that if $\alpha = 1 + \varepsilon$ where $\varepsilon \geq 0$, then

$$f(a+b)f(a \cdot b) = f(a)f(b) - [f(a) - f(a \cdot b)][f(b) - f(a \cdot b)] - \varepsilon[f(a \cdot b)]^2 \tag{19}$$

and also

$$\alpha.f(a+b)f(a\cdot b)=f(a)f(b)-[f(a)-\alpha f(a\cdot b)][f(b)-\alpha f(a\cdot b)]. \quad (20)$$

These immediately imply the following In-Ex inequalities.

Corollary 7.1. *(I) If $f(a\cdot b)\le f(a)$ then*

$$f(a+b)f(a\cdot b)\le f(a)f(b)-\varepsilon[f(a\cdot b)]^2\le f(a)f(b)\,. \quad (21)$$

If $\alpha=1$ then

$$f(a+b)f(a\cdot b)=f(a)f(b) \text{ iff } f(a)=f(a\cdot b)=f(b).$$

If $\alpha>1$ then

$$f(a+b)f(a\cdot b)=f(a)f(b) \text{ iff } f(a)=f(a\cdot b)=f(b)=0.$$

(II) If $\alpha f(a\cdot b)\le f(a)$ with $\alpha>1$, then

$$\alpha f(a+b)f(a\cdot b)=f(a)f(b) \text{ iff } f(a)=\alpha f(a\cdot b)=f(b)\,.$$

Corollary 7.2. *(i) If $f(a\cdot b)\le f(a)$, then*

$$f(a+b)f(a\cdot b)\le f(a)f(b)-\varepsilon[f(a\cdot b)]^2\le f(a)f(b)\,. \quad (22)$$

(ii) If $\alpha f(a\cdot b)\le f(a)$ then

$$f(a+b)f(a\cdot b)\le \frac{1}{\alpha}f(a)f(b)\,. \quad (23)$$

Needless to say, when $\alpha\ge 1$, the assumption that $\alpha f(a\cdot b)\le f(a)$ is much stronger than that of $f(a\cdot b)\le f(a)$. In particular, for XOR operation, this will generally NOT be true.

Examples with $\alpha=1$ and $\varepsilon=0$

(i) If $(S,+,\cdot)=(F(X),\cup,\cap)$ and $f(\cdot)=\#(\cdot)$, then we have

$$\#(A\cup B).\#(A\cap B)=\#(A).\#(B)-\#(A^c\cap B)\#(A\cap B^c)\le\#(A).\#(B) \quad (24)$$

(ii) If $S=\mathcal{F}$, a σ-field in $\mathcal{P}(X)$, $(+,\cdot)=(\cup,\cap)$ and $f(\cdot)=P(\cdot)$ a probability measure, then

$$P(A\cup B)P(A\cap B)=P(A)P(B)-P(A^c\cap B)P(A\cap B^c)\le P(A)P(B) \quad (25)$$

(iii) If $S = \mathbb{R}, (+, \cdot) = (\max, \min)$ and $f =$ identity, then

$$\max(a, b) \cdot \min(a, b)) = a \cdot b - (a - \min\{a, b\})(b - \min\{a, b\}) = ab \quad (26)$$

The latter follows from the fact that $a - \min\{a, b\} = \max\{a - b, 0\}$, therefore we get to $[a - min\{a, b\}][b - min\{a, b\}] = 0$.

(iv) Applying (iii) to integer prime powers we arrive at

$$[a, b](a, b) = a \cdot b.$$

As such we actually have equality in the latter two "inequalities".

(v) Let S be the collection of all subspaces of a vector space U, with $V + W$ being the vector space direct sum and $V \cdot W = V \cap W$. Then $f(V) = dim(V)$ is a valuation on S, and

$$\begin{aligned} dim(V_1 + V_2) dim(V_1 \cap V_2) &= dim(V_1) dim(V_2) - [dim(V_1) - dim(V_1 \cap V_2)] \\ &\quad \times [dim(V_1) - dim(V_1 \cap V_2)] \\ &\leq dim(V_1) dim(V_2). \end{aligned}$$

8 InEx equalities and inequalities for $n \leq 3$.

Consider InEx equation for three elements

$$f(a + b + c) = f(a) + f(b) + f(c) - \alpha[f(ab) + f(ac) + f(bc)] + \alpha^2 f(abc).$$

There are numerous equalities that follow from it, and some of these are non-trivial. Also the difference between the cases where $\alpha = 2$ (XOR operation) and $\alpha \neq 2$ becomes striking.

8.1 InEx equalities

We first present some of necessary equalities that must hold as a consequence of InEx for an α-valuation with $\alpha \geq 1$.

1. $f(ab + a) = f(a) - (\alpha - 1) f(ab) \geq f(a)$,
   which we refer to as a "weak version" of InEx. This in turn implies

2. $f(a + a) = (2 - \alpha) f(a)$ and thus $f(a + a) = f(a)$ iff $\alpha = 1$.
   For the XOR operation, this means that $f(a + a) = 0$.

3. $f(a + b + b) = f(a) + (2 - \alpha) f(b) - \alpha(2 - \alpha) f(ab)$.

4. We also have

$$\begin{aligned} f(ab+ac) &= f(ab+ac+bc) + \alpha(2-\alpha)t - f(bc) \\ &= f(ab+c) + f(ac) - f(c)\,. \end{aligned} \tag{27}$$

where $t = f(abc)$. Now if $\alpha f(ab) \le f(b)$ for all a and b then $\alpha t \le f(ab)$ as well as $\alpha t \le f(ac)$. We then have $f(ab+ac) = f(ab) + f(ac) - \alpha t \ge \alpha t + \alpha t - \alpha t = \alpha t$. In particular, for $\alpha \ge 1$,

$$f(abc) \le \alpha f(abc) \le f(ab+ac).$$

5. The following identities follow from InEx and can actually be used to characterize it for an α-valuation.

$$\begin{aligned} f(a) - f(ab+ac) &= f(a+c) - f(ab+c) + (\alpha-1)f(ac) \\ &= f(a) + f(c) - f(ac) - f(ab+c). \end{aligned} \tag{28}$$

Lemma 8.1. *The following are equivalent for $\alpha \ge 1$ and $n = 3$:*

(i) *InEx identity*
(ii) $f(a) - f(ab+ac) = f(a+c) - f(ab+c) + (\alpha-1)f(ac)$ *and* $f(ab+a) = (1-\alpha)f(ab) + f(a)$

(29)

Proof. The necessity follows from the above equalities (28). For sufficiency, suppose (29) holds, and set $c = b$. Then we get $f(a) - f(ab+ab) = f(a+b) - f(ab+b) + (\alpha-1)f(ab)$ in which we substitute the weak InEx identity $f(ab+b) = (1-\alpha)f(ab) + f(b)$ to give the desired result. $\square$

8.2 InEx Inequalities

We next come to some of the inequalities spawned by InEx. First, we note that $f(a+b) - f(a) = f(b) - \alpha f(ab)$, with $\alpha \ge 1$, ensures that

$$\alpha f(ab) \le f(b) \Leftrightarrow f(a) \le f(a+b). \tag{30}$$

A useful consequence is

Corollary 8.1.

$$t = f(abc) \le \alpha f(abc) \le f(ab) \le f(ab+ac). \tag{31}$$

We also have

Lemma 8.2. *Let* $\alpha f(ab) \leq f(a)$ *with* $\alpha \geq 1$*. Then all the following statements hold.*

(i) $f(ab) \leq f(a)$ *for all* a *and* b*,*

(ii) $f(a(b+c)) \leq f(a)$ *for all* a*,*b *and* c*,*

(iii) $f(ab+c) \leq f(b)+f(c)-f(bc)$ *for all* a*,*b *and* c*,*

(iv) $f(ab+c) \leq f(b+c)+\varepsilon f(bc)$ *for all* a*,*b *and* c*,*

(v) $f(ab)-\alpha f(abc) \leq f(b)-f(bc)$ *for all* a*,*b *and* c *(monotonicity inequality)*

Proof. (i) holds and also clearly (i) implies (ii). The equivalence of (ii), (iii) and (iv) follows at once from the identity:

$$0 \leq f(b)-f(ab+bc) = f(b+c)-f(ab+c)+\varepsilon f(bc) = f(b)+f(c)-f(bc)-f(ab+c).$$

The equivalence of (ii) and (v) follows from the identity

$$0 \leq f(b)-f(ab+bc) = f(b)-f(ab)-f(bc)+\alpha f(abc).$$

□

For $\alpha = 1$, we may combine Lemma 8.2 (v) with the inequality (30). In addition, the monotone inequality corresponds to $q \leq q+r$ in the following Venn diagram.

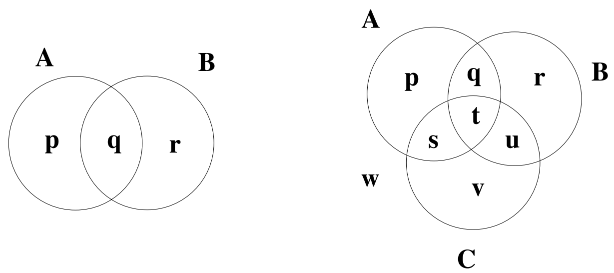

Figure 1: Three sets

Furthermore, it is clear from (30) that

$$0 \leq \alpha^2 f(abc) \leq \alpha f(ab) \leq f(a)$$

and

$$3\alpha^2 f(abc) \leq \alpha[f(ab)+f(ac)+f(bc)] \leq [f(a)+f(b)+f(c)]\,.$$

8.3 Generalizations of the InEx Inequality to 3-d.

Let us now examine possible generalization of the 2-d InEx inequality

$$f(a+b)f(ab) \leq f(a)f(b)\,, \tag{32}$$

to the sum $f(a+b+c)$, assuming that $\alpha = 1$ and $f(ab) \leq f(a)$.

There are numerous "sharp" generalizations, that all reduce to (32), when two of the variables are set equal. We shall present three of these under the assumption that the addition is also idempotent.

Our first method is to replace b by $b+c$ in the inequality (32). This gives

$$f(a+b+c)f[a(b+c)] \leq f(a)f(b+c), \tag{33}$$

which we combine with

$$f(abc) \leq f(ab) \leq f(ab+ac) \tag{34}$$

to arrive at

$$f(a+b+c)f(abc) \leq f(a+b+c)f[a(b+c)] \leq f(a)f(b+c)\,. \tag{35}$$

Rotating the variables we may conclude that

$$f(a+b+c)f(abc) \leq \min\{f(a)f(b+c), f(b)f(a+c), f(c)f(a+b)\}. \tag{36}$$

This result is sharp, in the sense that if we set $b = c$, this reduces to

$$f(a+b)f(ab) \leq \min\{f(a)f(b), f(b)f(a+b), f(b)f(a+b)\} = f(a)f(b)\,,$$

as $f(a) \leq f(a+b)$.

Alternatively, we may multiply (35) by f(bc) and use InEx identity for $n = 2$, to arrive at

$$f(a+b+c)f(abc)f(bc) \leq f(a)f(b+c)f(bc) \leq f(a)f(b)f(c)\,. \tag{37}$$

Rotating the variables then yields

$$f(a+b+c)f(abc)\max\{f(ab), f(ac), f(bc)\} \leq f(a)f(b)f(c)\,. \tag{38}$$

Setting b = c, shows that $f(a+b)f(ab)\max\{f(ab), f(ab), f(b)\} \leq f(a)f(b)f(b)$, which reduces to (32). Based on the inequalities that we have seen for the additive case, it would be natural to expect that for a sub-multiplicative valuation, with $f(ab) \leq f(a)$ and one may expect the following to hold

$$f(a+b+c)f(ab)f(ac)f(bc) \leq f(a)f(b)f(c)f(abc). \tag{39}$$

However, numerical tests prove that this is not true, even for probability measures. In fact, the inequality (39)

$$\begin{aligned} &(t+p+q+r+s+u+v)(t+q)(t+s)(t+u) \leq \\ &\leq t(t+p+q+s)(t+q+r+u)(t+s+u+v) \end{aligned} \tag{40}$$

is violated when we set $p = q = r = s = u = v = \frac{1}{10}$ and $t = \frac{1}{100}$ for instance.

However, a slight perturbation from this does hold as below.

$$\begin{aligned} &(t+p+q+r+s+u+v)[(t+q)(t+s)(t+u) - qsu] \leq \\ &\leq t[(t+p+q+s)(t+q+r+u)(t+s+u+v) + qsu]. \end{aligned} \tag{41}$$

We may re-express this as

$$f(a+b+c)[f(ab)f(ac)f(bc) - qsu] \leq f(abc)[f(a)f(b)f(c) + qsu], \tag{42}$$

where $q = f(ab) - f(abc), s = f(ac) - f(abc)$ and $u = f(bc) - f(abc)$. If we set $b = c$, then $q = 0$, and we are back to $n = 2$ case.

Note that the outer layer of "petals" are p, r and v while the inner layer is made up of q, s and u. The sum $ab + ac + bc$ corresponds to the inner "flower" of the Venn diagram, made up of the petals q, s, u and t. The proof of the inequality (42) will be given later when we actually clarify when the equality holds.

Examples

(i) For a probability measure, we may state:

$$P(A \cup B)P(A \cap B) \leq P(A)P(B) \tag{43}$$

with the equality holding if and only if $P(A \setminus B) = 0 = P(B \setminus A)$. Also we have

$$P(A \cup B \cup C)P(A \cap B \cap C) \leq \min\{P(A)P(B \cup C), P(B)P(A \cup B), P(C)P(A \cup B)\} \tag{44}$$

as well as

$$P(A \cup B \cup C)P(A \cap B \cap C) \max\{P(A \cap B), P(A \cap C), P(B \cap C)\} \leq P(A)P(B)P(C). \tag{45}$$

(ii) For a probability measure, the inequality (42) becomes

$$\begin{aligned} P(A\cup B\cup C)\Big[&P(A\cap B)P(A\cap C)P(B\cap C)\\ -\Big[&[P(A\cap B)-P(A\cap B\cap C)][[P(A\cap C)-P(A\cap B\cap C)]\\ &\times[P(B\cap C)-P(A\cap B\cap C)]\Big]\Big] \le P(A\cap B\cap C)\Big[P(A)P(B)P(C)\\ +&[P(A\cap B)-P(A\cup B\cup C)][P(A\cap C)-P(A\cap B\cap C)]\\ &\times[P(B\cap C)-P(A\cap B\cap C)]\Big], \end{aligned} \tag{46}$$

which reduces to (43) when $B=C$.

To see when we actually achieve equality in (42), we start by expressing both sides as polynomials in t. We let

$$\begin{aligned}(t+p+q+r+s+u+v)&[(t+q)(t+s)(t+u)-qsu]\\ &=t^4+e_3t^3+e_2t^2+e_1t+e_0-f(a+b+c)\cdot qsu\end{aligned}$$

and set

$$t[(t+p+q+s)(t+q+r+u)(t+s+u+v)+qsu]=t^4+f_3t^3+f_2t^2+f_1t+t.qsu\,.$$

For the two sides to be equal we must have

$$(f_3-e_3)t^3+(f_2-e_2)t^2+(f_1-e_1)t+qsu.t+f(a+b+c).qsu-e_0=0. \tag{47}$$

Now note that $f_3=e_3$ while $f_2-e_2=(pu+qv+rs)+(pr+pv+rv)$, $e_0=[f(a+b+c)-t](qsu)$ and $f_1-e_1=\lambda-2qsu$, where

$$\lambda=(pqv+pvu)+(prs+pru)+(rsv+rvq)+(q^2v+u^2p+s^2r)+prv.$$

Substituting these gives

$$\begin{aligned}&[(pu+qv+rs)+(pr+pv+rv)]t^2+(\lambda-2qsu)t+(qsu)t+\\ &+f(a+b+c)qsu-[f(a+b+c)-t]qsu=0\end{aligned}$$

and consequently

$$t[t(f_2-e_2)+\lambda]=0.$$

We can now conclude that the inequality (42) must hold since the difference between the two sides in (41) is given by $t[t(f_2-e_2)+\lambda]$, in which each term is non-negative.

We close by examining the equality case. Indeed, since all terms are non-negative, they must vanish and we have the following two cases.

Case (i) $t = 0$ or Case (ii) $t \neq 0$. In the latter case we must have $f_2 = e_2$ and $\lambda = 0$.

The equality $f_2 = e_2$ ensures that

$$pu = 0, qv = 0, rs = 0, pr = 0, pv = 0, rv = 0. \tag{48}$$

The first three contain "cross products" between inner and outer petals while the latter three involve only the three outer petals. These conditions in turn imply that $\lambda = 0$.

Let us close with some relevant comments and open questions.

9 Remarks and open Questions

1. We have seen a partial parallel between the additive and multiplicative inequalities for valuations. It would be of interest to find more general multiplicative inequalitiesfor valuations.
2. In how far does convexity play a role? We have met the consequences of composing an evaluation map with the exponential functions. If would be interesting to explore the composition of an evaluation map with other convex or concave functions.
3. Investigations into **sub**-valuation for which $f(a + b) \leq f(a) + f(b) - \alpha f(a \cdot b)$ would be of interest. The catch, however, is that inequalities cannot be lined up in this case.
4. Do analogous inequalities exist for multinomial coefficients?
5. The derivation of the valuation inequalities may be done by replacing the assumption that an additive inverse exists, by the assumption that a partial order exists.

References

[1] G.Pólya and G. Szegö, *Aufgabe und Lehrsätze aus der Analysis*, Springer Berlin, 1964, 3rd ed, pp. 119-121.

[2] D.E. Rutherford, *Introduction to Lattice Theory*, Hafner, New York, 1965, p. 20.

[3] G. Birkhoff, *Lattice theory*, vol 25, 3rd ed, Providence RI, AMS series.