Unified Approach to Probability Problems and Estimation Algorithms Associated With Symmetric Functions

David Hertz

Correspondence: Independent Researcher, Akko, Israel. E-mail: hertz@013.net.il or drdavidhertz@gmail.com

Received: April 19, 2021 Accepted: June 25, 2021 Online Published: June 28, 2021

doi:10.5539/ijsp.v10n4p166 URL: https://doi.org/10.5539/ijsp.v10n4p166

Abstract

In this article we introduce a simple powerful methodology where we replace the independent variables $\lambda_1, ..., \lambda_n$ in various symmetric functions as well as in Vieta's formulas by the indication functions of the events $A_i, i = 1, ..., n$, i.e., $\lambda_i = \mathbf{1}(A_i), i = 1, ..., n$. Both the random variable K that counts the number of events that actually occurred and the proposed obvious identity $\Pi_{i=1}^n (z - \mathbf{1}(A_i)) \equiv (z-1)^K z^{n-K}$ that solely depends on K play a central role in this article. Just by choosing different values for z (real, complex, and random) and taking expectations of the various functions we provide other simple proofs of known results as well as obtain new results. The estimation algorithms for computing the expected elementary symmetric functions via least squares based on **IFFT** in the complex domain ($z \in \mathbb{C}$) and least squares or linear programming in the real domain ($z \in \mathbb{R}$) are noteworthy. Similarly, we we use Newton's identities and some well known inequalities to obtain new results and inequalities. Then, we give an algorithm that exactly computes the distribution of K (i.e., $q_k := \mathbb{P}(K = k), k = 0, 1, ..., n$) for finite sample spaces. Finally, we give the conclusion and area for further research.

1. Introduction

This article is self contained and the results appear in the order that they were conceived. Let

$$f_n(z) := z^n + a_1 z^{n-1} + \cdots + a_n \tag{1}$$

be a real monic polynomial and $\lambda_1, \lambda_2, ..., \lambda_n$ its roots. We have

$$\begin{aligned} f_n(z) &= \textstyle\prod_{j=1}^n (z - \lambda_j) \\ &= z^n - e_1(\lambda_1, \ldots, \lambda_n) z^{n-1} + e_2(\lambda_1, \ldots, \lambda_n) z^{n-2} + \cdots + (-1)^n e_n(\lambda_1, \ldots, \lambda_n), \end{aligned} \tag{2}$$

where $e_i := (-1)^i a_i,\ i = 1, ..., n$ are the elementary symmetric functions. Since $f_n(z)$ is monic we define $e_0 := 1\ \forall n$. These relations between the roots and the coefficients of a polynomial are called Vieta's formulas.

The elementary symmetric functions for $n = 1, 2, 3$ are (Wiki contributions, 2019 Oct. 19)

For $n = 1$:

$$e_1(\lambda_1) = \lambda_1. \tag{3}$$

For $n = 2$:

$$\begin{aligned} e_1(\lambda_1, \lambda_2) &= \lambda_1 + \lambda_2 \\ e_2(\lambda_1, \lambda_2) &= \lambda_1 \lambda_2. \end{aligned} \tag{4}$$

For $n = 3$:

$$\begin{aligned} e_1(\lambda_1, \lambda_2 \lambda_3) &= \lambda_1 + \lambda_2 + \lambda_3 \\ e_2(\lambda_1, \lambda_2, \lambda_3) &= \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 \\ e_3(\lambda_1, \lambda_2, \lambda_3) &= \lambda_1 \lambda_2 \lambda_3. \end{aligned} \tag{5}$$

These multi-variable polynomials are called symmetric since their value is preserved under permutations. For example, $e_2(\lambda_3, \lambda_2, \lambda_1) \equiv e_2(\lambda_1, \lambda_2, \lambda_3)$. Next, we will present brief formulas for $e_k, k = 1, .., n$. These formulas are based on the following definition (Marcus & Minc, 1964, p. 10).

Definition 1.1. *If $1 \le k \le n$, then $Q_{k,n}$ will denote the totality of strictly increasing sequences of k integers chosen from $1, ..., n$.*

Hence, using this notation we can write

$$e_k(\lambda_1, ..., \lambda_n) = \sum_{\alpha \in Q_{k,n}} \prod_{i=1}^{k} \lambda_{\alpha_i}, \tag{6}$$

where $\alpha := (\alpha_1, ..., \alpha_k)$ and $1 \le \alpha_1 < \cdots < \alpha_k \le n$. Notice that the cardinality of $Q_{k,n}$ is

$$|Q_{k,n}| = \binom{n}{k}. \tag{7}$$

Other symmetric polynomials, to name but a few, are:

(i) Power sums:

$$p_k(\lambda_1, ..., \lambda_n) := \sum_{i=1}^{n} \lambda_i^k, \quad k = 1, 2, ... \tag{8}$$

where $p_1 \equiv e_1$ and the other $p_k, k > 1$ are related to $e_i, i \le k$ via Newton's identities, see (Wikipedia contributions, 2021, Feb. 4).

(ii) The sum of all monomials of total degree k which are called complete homogeneous symmetric functions and are denoted by $h_k, k = 1, ..., n$. Next, we will present brief formulas for the h_k's. These formulas are based on the following definition (Marcus & Minc, 1964, p. 10).

Definition 1.2. *If $1 \le k \le n$, then $G_{k,n}$ will denote the totality of non-decreasing sequences of k integers chosen from $1, ..., n$.*

Hence, using this notation we can write

$$h_k(\lambda_1, ..., \lambda_n) = \sum_{\alpha \in G_{k,n}} \prod_{i=1}^{k} \lambda_i^{\alpha_i}, \tag{9}$$

where $\alpha := (\alpha_1, ..., \alpha_k)$ and $1 \le \alpha_1 \le \cdots \le \alpha_k \le n$. Notice that the cardinality of $G_{k,n}$ is

$$|G_{k,n}| = \binom{n+k-1}{k-1} = \binom{n+k-1}{n}. \tag{10}$$

In what follows we let

$$\lambda_i = \mathbf{1}(A_i), \;\; i = 1, ..., n. \tag{11}$$

Therefore,

$$K := e_1 = \mathbf{1}(A_1) + \cdots + \mathbf{1}(A_n) \tag{12}$$

is a random variable indicating how many of the events $A_i, i = 1, ..., n$ actually occurred. Hence, substituting $\lambda_i = \mathbf{1}(A_i), i = 1, ..., n$ in (2) we obtain

$$f_n(z) = (z - \mathbf{1}(A_1)) \cdots (z - \mathbf{1}(A_n)). \tag{13}$$

Obviously,

$$f_n(z) = (z-1)^K z^{n-K}. \tag{14}$$

Equations (11) - (14) are the key equations to derive all the results that follow. When $K = 0$, combining (13) and (14) and noting that $\mathbf{1}(A_i) = 0, i = 1, ..., n$ we obtain $z^n \equiv z^n(z-1)^0$ for all z. In particular, substituting $z = 1$ we are led to define

$$0^0 := 1. \tag{15}$$

Using (13) and Vieta's formulas we obtain

$$f_n(z) := z^n + (-1)^1 e_1 z^{n-1} + (-1)^2 e_2 z^{n-2} + ... + (-1)^n e_n, \tag{16}$$

where

$$\begin{aligned} e_k &= e_k(\mathbf{1}(A_1), ..., \mathbf{1}(A_n)) \\ &= \sum_{\alpha \in Q_{k,n}} \prod_{i=1}^{k} \mathbf{1}(A_{\alpha_i}) \\ &= \sum_{\alpha \in Q_{k,n}} \mathbf{1}\left(\cap_{i=1}^{k} A_{\alpha_i}\right). \end{aligned} \tag{17}$$

For now we assume that $z \in \mathbb{C}$ is a constant, later on we will let z be a random variable, say Z. Assuming that $z \neq 0$, using (13) - (17), and taking expectations, we obtain

$$z^n \mathbb{E}\left(1 - z^{-1}\right)^K = z^n + (-1)^1 \mathbb{E} K z^{n-1} + \ldots + (-1)^k \sum_{\alpha \in Q_{k,n}} \mathbb{E}\mathbf{1}\left(\cap_{i=1}^{k} A_{\alpha_i}\right) z^{n-k} + \ldots + (-1)^n \mathbb{E}\mathbf{1}\left(\cap_{i=1}^{n} A_i\right). \tag{18}$$

Notice that

$$\mathbb{E}\mathbf{1}(A) = 1\mathbb{P}(A) + 0(1 - \mathbb{P}(A)) = \mathbb{P}(A), \ \forall A \subset \Omega, \tag{19}$$

where Ω denotes the sampling space. Hence, we have

$$z^n \mathbb{E}\left(1 - z^{-1}\right)^K = \sum_{k=0}^{n} (-1)^k \overline{e}_k z^{n-k}, \ z \neq 0, \tag{20}$$

where

$$\overline{e}_k = \sum_{\alpha \in Q_{k,n}} \mathbb{P}(\cap_{i=1}^{k} A_{\alpha_i}), k = 1, ..., n. \tag{21}$$

The organization of this paper is as follows. In Section 2 we will deal with various values of z. In Subsection 2.1 we let $z = 0$ and obtain that $\mathbb{P}(K = n) = \mathbb{P}(\cap_{i=1}^{n} A_i)$. In Subsection 2.2 we let $z = 1$ and obtain a simple proof of the inclusion exclusion formula. In Subsection 2.3 we let $z = -1$ and obtain an explicit expression for $\mathbb{E}(2^K)$, i.e., the expected number of subfamilies of $\{A_1, ..., A_n\}$ all whose component events occur. In Subsection 2.4 we let $z^{-1} = 1 - s$, $z^{-1} = 1 - e^s$, and $z^{-1} = 1 - e^{it}$ take expectations and thus obtain K's probability generating function, $G_K := \mathbb{E}(s^K)$, moment generating function, $M_K(s) := \mathbb{E}(e^{sK})$, and characteristic function, $\phi_K(t) := \mathbb{E}(e^{itK})$. Using the probability generating function we present a simpler proof of Waring's formulas that relates the q_k's to the $\overline{e}_k$'s. In Subsection 2.5 we let $z_k = \exp(\frac{2\pi i}{N} k), k = 0, 1, ..., N-1, N \geq n, i = \sqrt{-1}$, and present an efficient least squares algorithm based on an N-point **IFFT** (Inverse Fast Fourier Transform) to estimate the $\overline{e}_k$'s. We will prove that if $N > n$ the proposed algorithm becomes a least squares solution. In Subsection 2.6 we choose distinct $z_k \in \mathbb{R} \setminus 0, k = 1, ..., N, N \geq n$ and give algorithms based on least squares (LS) and linear programming (LP) to estimate the elementary symmetric functions. In Subsection 2.7 we let $z = Z$ be a random variable and give two examples. In the first example we let $Z = \mathbf{1}(A_1)$ and obtain two interesting identities. In the second example we let $Z \sim N(0, \sigma^2)$, assume that Z is independent of the A_i's, and thus arrive at quite complex formulas for the expected values. In Section 3 we present formal proofs of known formulas that relate $\overline{e}_k := \mathbb{E} e_k, k = 0, 1, ..., n$ to $q_k := \mathbb{P}(K = k), k = 0, 1, ..., n$ and give some consequences. In Section 4 We present Newton's identities and give some consequences. In Section 5 we present several inequalities and a conjecture.

In Section 6 we present a polynomial time algorithm to compute K's probability density function, i.e., the q_k's. Using these q_k's and the results of Section 3 we compute the $\overline{e}_k$'s for a finite sample space. Finally, in Section 7 we give the conclusion and mention topics for further research.

In what follows we will use capital letters to denote random variable and boldface letters to denote vectors and matrices. Notice that here the indices of all vectors and matrices start at one.

2. Special Cases for Various Values of z Consequences and Algorithms

2.1 The Case $z = 0$.

Using (14) and (2) the case $z = 0$ gives the rather trivial identities

$$\mathbb{E}((-1)^K 0^{n-K}) = (-1)^n \mathbb{E}\mathbf{1}(K = n) \equiv (-1)^n \mathbb{E}\mathbf{1}\left(\cap_{i=1}^n A_i\right). \tag{22}$$

Hence, using (19) we obtain

$$\mathbb{E}\mathbf{1}(K = n) = \mathbb{E}\mathbf{1}\left(\cap_{i=1}^n A_i\right) = \mathbb{P}\left(\cap_{i=1}^n A_i\right). \tag{23}$$

Using Demorgan's law we also obtain

$$\mathbb{E}\mathbf{1}(K = 0) \equiv \mathbb{E}\mathbf{1}\left(\cap_{i=1}^n A_i^c\right) = \mathbb{E}\mathbf{1}\left((\cup_{i=1}^n A_i)^c\right) = \mathbb{P}\left((\cup_{i=1}^n A_i)^c\right) = 1 - \mathbb{P}\left(\cup_{i=1}^n A_i\right), \tag{24}$$

where A_i^c denotes the complement of A_i.

2.2 The Case $z = 1$.

When $z = 1$ we provide a simple proof of the following known probability rule, called the inclusion exclusion formula.

Theorem 2.1.

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i<j}^n \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k}^n \mathbb{P}(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n+1}\mathbb{P}(A_1 \cap \cdots \cap A_n). \tag{25}$$

By using (21) this equation can be written in shorter notation as follows.

$$\begin{aligned}\mathbb{P}(\cup_{i=1}^n A_i) &= \sum_{k=1}^n \left((-1)^{k+1} \sum_{\alpha \in Q_{k,n}} \mathbb{P}(\cap_{i=1}^k A_{\alpha_i}) \right) \\ &= \sum_{k=1}^n (-1)^{k+1} \overline{e}_k.\end{aligned} \tag{26}$$

Proof: Substituting in (18) $z = 1$ and using (19) we obtain

$$\mathbb{E}0^K = 1 + (-1)^1 \sum_{i=1}^n \mathbb{P}(A_i) + \ldots + (-1)^k \sum_{\alpha \in Q_{k,n}} \mathbb{P}\left(\cap_{i=1}^k A_{\alpha_i}\right) + \ldots + (-1)^n \mathbb{P}\left(\cap_{i=1}^n A_i\right). \tag{27}$$

Since, $0^0 = 1$, we obtain that $\mathbb{E}0^K = \mathbb{E}\mathbf{1}(K = 0)$. Finally, by using (24) and some algebra we obtain the desired result. ■

Notice that a similar derivation appears in (Grimmet & Stirzaker, 2001, p. 56), where the authors by nontrivial manipulations of indicator functions obtained

$$\mathbf{1}(\cup_{i=1}^n A_i) = 1 - \prod_{i=1}^n (1 - \mathbf{1}(A_i)). \tag{28}$$

Then, by using Vieta's formulas and taking expectations they arrived at (25).

2.3 The Case $z = -1$.

When $z = -1$ we provide a simple proof of the following nontrivial probability rule.

Theorem 2.2. *Let $q_k := \mathbb{P}(K = k)$, then*

$$\mathbb{E}(2^K) = \sum_{k=0}^n q_k 2^k = 1 + \sum_{k=1}^n \overline{e}_k, \tag{29}$$

where the $\overline{e}_k$'s are given by (21). Notice that $\mathbb{E}(2^K)$ is the expected number of subfamilies of $\{A_1, ..., A_n\}$ all of whose component events occur.

Proof: The first equation follows from the definition of $\mathbb{E}(2^K)$. The second equation follows by substituting $z = -1$ in (20) and carrying out some algebra. ■

In Section 3 we will give another simple proof of this formula.

2.4 The Cases $z^{-1} = 1 - s$, $z^{-1} = 1 - e^s$, and $z^{-1} = 1 - e^{it}$.

In this subsection we will show how to obtain expressions for K's probability generating function, $G_K := \mathbb{E}(s^K)$, moment generating function, $M_K(s) := \mathbb{E}(e^{sK})$, and characteristic function, $\phi_K(t) := \mathbb{E}(e^{itK})$ by appropriately choosing $z \neq 0$. Using (20) and substituting $z = y^{-1}$ we obtain

$$(1 - y)^K = 1 - \sum_{k=1}^{n} (-1)^k e_k y^k. \tag{30}$$

Hence, substituting $y = 1 - s$ and taking expectations we obtain the probability generating function of K [1, p. 150], i.e.,

$$G_K(s) := \mathbb{E}(s^K) = \sum_{k=0}^{n} (-1)^k \overline{e}_k (1 - s)^k. \tag{31}$$

We also have

$$G_K(s) := \sum_{k=0}^{n} q_k s^k. \tag{32}$$

In the next theorem by equating the coefficients of the last two equations we will obtain a simpler straightforward proof to the following Waring's formulas (Grimmet & Stirzaker, 2001, p.151).

Theorem 2.3. *Let $q_k := \mathbb{P}(K = k)$, then*

$$q_\ell = \sum_{k=\ell}^{n} (-1)^{k+\ell} \binom{k}{\ell} \overline{e}_\ell, \ell = 0, ..., n, \tag{33}$$

where the $\overline{e}_k$'s are given by (21).

Proof:

$$\begin{aligned} G_K(s) &= \sum_{k=0}^{n} q_k s^k \\ &= \sum_{k=0}^{n} (-1)^k \overline{e}_k (1 - s)^k \\ &= \sum_{k=0}^{n} (-1)^k \overline{e}_k \left(\sum_{\ell=0}^{k} (-1)^\ell \binom{k}{\ell} s^\ell \right) \\ &= \sum_{\ell=0}^{n} s^\ell \left(\sum_{k=\ell}^{n} (-1)^{k+\ell} \binom{k}{\ell} \right) \\ &= \sum_{k=0}^{n} s^k \left(\sum_{\ell=k}^{n} (-1)^{k+\ell} \binom{\ell}{k} \right). \end{aligned} \tag{34}$$

Equating coefficients of s^k we obtain the desired result. This completes the proof. ■

Next, substituting in (30), $y = 1 - e^s$ and taking expectations we obtain the moment generating function of K (Grimmet & Stirzaker, 2001, p. 151):

$$M_K(s) := \mathbb{E}(e^{sK}) = \sum_{k=0}^{n} (-1)^k \overline{e}_k (1 - e^s)^k. \tag{35}$$

Finally, substituting in (30), $y = 1 - e^{it}$ and taking expectations we obtain the characteristic function of K (Grimmet & Stirzaker, 2001, p. 182) :

$$\phi_K(t) := \mathbb{E}(e^{itK}) = \sum_{k=0}^{n} (-1)^k \overline{e}_k (1 - e^{it})^k. \tag{36}$$

2.5 The Case $z_k = \exp(\frac{2\pi i}{N}k), k = 0, 1, ...N-1, N \geq n,\ i = \sqrt{-1}$ and Efficient Algorithm for Estimating the $\overline{e}_k$'s.

Here, we present an efficient algorithm based on an N-point **IFFT** to estimate the $\overline{e}_k$'s. We will prove that if $N > n$ the proposed algoritm gives a least squares solution. Substituting in (20) $z_k = \exp(\frac{2\pi i}{N}k), k = 0, 1, ..., N-1$ we obtain a set of N equations in the n unknown alternating expected elementary symmetric functions, i.e.,

$$\mathbf{X} := (x_0, ..., x_{n-1})^T := ((-1)^1\overline{e}_1, (-1)^2\overline{e}_2, ..., (-1)^n\overline{e}_n)^T. \tag{37}$$

Since $\overline{e}_0 = 1$, $z_k^N = 1$, $z_k^{-1} = z_{N-k}$ for $k = 0, ..., N-1$, we obtain

$$\begin{aligned} z_k^n\mathbb{E}\left((1 - z_{N-k})^K\right) - z_k^n &= \sum_{\ell=1}^{n}(-1)^\ell\overline{e}_\ell \exp\left(\frac{2\pi i}{N}(n-\ell)k\right) \\ &= z_k^n\sum_{\ell=1}^{n}(-1)^\ell\overline{e}_\ell \exp\left(-\frac{2\pi i}{N}\ell k\right) \\ &= z_k^{n-1}\sum_{\ell=0}^{n-1}(-1)^{\ell+1}\overline{e}_{\ell+1} \exp\left(-\frac{2\pi i}{N}\ell k\right). \end{aligned} \tag{38}$$

Hence, dividing both sides by z_k^{n-1} we obtain

$$\begin{aligned} z_k\left(\mathbb{E}(1 - z_{N-k})^K - 1\right) &= \sum_{\ell=0}^{n-1}(-1)^{\ell+1}\overline{e}_{\ell+1} \exp\left(-\frac{2\pi i}{N}\ell k\right) \\ &= \sum_{\ell=0}^{n-1} x_\ell \exp\left(-\frac{2\pi i}{N}\ell k\right). \end{aligned} \tag{39}$$

Let

$$\mathbf{Y} := (y_0, y_1, ..., y_{N-1})^T,$$

where

$$y_k := z_k(\mathbb{E}(1 - z_{N-k})^K - 1), k = 0, ..., N-1. \tag{40}$$

Also let

$$\mathbf{F} := [\mathbf{F}_{k,\ell}] = \left[\exp\left(-\frac{2\pi i}{N}\ell k\right), k, \ell = 0, ..., N-1\right]. \tag{41}$$

We have

$$\mathbf{F}^{-1} = \frac{1}{N}\mathbf{F}^*, \tag{42}$$

where * denotes the hermitian operator.

Equation (39) can be written in matrix form as follows.

$$\begin{aligned} \mathbf{Y} &= \mathbf{F}_1\mathbf{X} \\ &= \mathbf{F}_1\mathbf{X} + \mathbf{F}_2\mathbf{0}_{N-n} \\ &= (\mathbf{F}_1, \mathbf{F}_2)\begin{pmatrix}\mathbf{X} \\ \mathbf{0}_{N-n}\end{pmatrix} \\ &= \mathbf{F}\begin{pmatrix}\mathbf{X} \\ \mathbf{0}_{N-n}\end{pmatrix} \\ &= \mathbf{DFT}\begin{pmatrix}\mathbf{X} \\ \mathbf{0}_{N-n}\end{pmatrix}, \end{aligned} \tag{43}$$

where $\mathbf{0}_\ell$ is a column vector of ℓ zeros,

$\mathbf{F}_1$ is the submatrix consisting of the first n columns of $\mathbf{F}$, and $\mathbf{F}_2$ is the submatrix consisting of the last $N-n$ columns of $\mathbf{F}$.

Usually, $\mathbf{Y}$ is not known. Hence, in what follows we will replace $\mathbf{Y}$ by an estimator, say $\hat{\mathbf{Y}}$. The least squares solution for given $\hat{\mathbf{Y}}$, say $\hat{\mathbf{X}}$, is given by

$$\begin{aligned}\hat{\mathbf{X}} &= (\mathbf{F}_1^*\mathbf{F}_1)^{-1}\mathbf{F}_1^*\hat{\mathbf{Y}} \\ &= \frac{1}{N}\mathbf{I}_n\mathbf{F}_1^*\hat{\mathbf{Y}} \\ &= \frac{1}{N}\mathbf{F}_1^*\hat{\mathbf{Y}} \\ &= [\mathbf{I}_n, \mathbf{0}_{n,N-n}]\mathbf{F}^{-1}\hat{\mathbf{Y}} \\ &= [\mathbf{I}_n, \mathbf{0}_{n,N-n}]\mathbf{IDFT}(\hat{\mathbf{Y}}),\end{aligned} \tag{44}$$

where **IDFT** denotes the inverse **DFT**. Hence, the least squares estimator $\hat{\mathbf{X}}$ of $\mathbf{X}$ can be computed by $\mathbf{IDFT}(\hat{\mathbf{Y}})$ and retaining its first n elements.

Remark 2.4. *Notice that since* $\mathbf{X} \in \mathbb{R}^n$ *we have*

$$\begin{aligned}y_k^* &= \sum_{\ell=0}^{n-1} x_\ell \exp\left(\frac{2\pi ik}{N}\ell\right) \\ &= \sum_{\ell=0}^{n-1} x_\ell \exp\left(-\frac{2\pi i(N-k)}{N}\ell\right) \\ &= y_{N-k},\ k = 1, ..., \left\lfloor \frac{N-1}{2} \right\rfloor,\end{aligned} \tag{45}$$

*where here * denotes conjugation. Hence,*

$$\hat{y}_{N-k} = \hat{y}_k^*,\ k = 1, ..., \left\lfloor \frac{N-1}{2} \right\rfloor. \tag{46}$$

Since $N - \lfloor (N-1)/2 \rfloor = \lceil N/2 \rceil + 1$ *we only need to compute* $\hat{y}_k, k = 0, ..., \lceil N/2 \rceil$ *and then use (46) to obtain* $\hat{\mathbf{Y}}$.

Hence, we propose the following algorithm to compute the least squares estimator $\hat{\mathbf{X}}$ of $\mathbf{X}$.

Algorithm 2.5.

$\hat{\mathbf{q}} = (\hat{q}_0, \hat{q}_1, ..., \hat{q}_n) = 0$
Let $\{\omega_1, \omega_2, ..., \omega_M\}$ be M random samples from Ω
for $m = 1, ..., M$
  $k = \sum_{\ell=1}^n \mathbf{1}(\omega_m \in A_\ell)$
  $\hat{q}_k = \hat{q}_k + 1$
end
$\hat{q} = \hat{q}/M$
$\hat{\mathbf{Y}} = (\hat{y}_0, ..., \hat{y}_{N-1})^T = 0$
for $k = 0, ..., \lceil N/2 \rceil$
for $\ell = 1, ..., n$
  $\hat{y}_k = \hat{y}_k + z_k\hat{q}_\ell\left((1 - z_{N-k})^\ell - 1\right)$
end
end
for $k = 1, ..., \lfloor (N-1)/2 \rfloor + 1$
  $\hat{y}_{N-k} = \hat{y}_k^*$
end
$\hat{\mathbf{X}} = (\mathbf{I}_n, \mathbf{0}_{n \times N-n})\mathbf{IFFT}(\hat{\mathbf{Y}})$.

Hence, using the definition of $\mathbf{X}$ in (37) we obtain the following estimator for $\overline{e}_k$, i.e.,

$$\hat{e}_k = |\hat{\mathbf{X}}(k)|, k = 1, ..., n. \tag{47}$$

Next, using (26) and (29) we obtain the following two estimators

$$\hat{\mathbb{P}}(\cup_{i=1}^{n} A_i) = -\sum_{i=1}^{n} \hat{\mathbf{X}}(i) \tag{48}$$

and

$$\hat{\mathbb{E}}(2^K) = 1 + \sum_{i=1}^{n} |\hat{\mathbf{X}}(i)|. \tag{49}$$

The complexity of the proposed algorithm is $O(Mn)$ to compute the $(n+1)$-dimensional vector $\hat{\mathbf{q}}$, $O(Nn)$ to compute the vector $\hat{\mathbf{Y}}$, and $O(N\log(N))$ to compute $\mathbf{IFFT}(\hat{\mathbf{Y}})$. Summing up, we obtain that the complexity of the proposed algorithm is $O(n\max\{N, M\})$. Notice that the proposed algorithm is applicable to finite as well as infinite sample spaces Ω and that for finite Ω the number of events n can be as high as $O(2^{|\Omega|})$.

2.6 The Case of Distinct $z_k \in \mathbb{R} \setminus 0, k = 1, ..., N, N \geq n$ and Algorithms

Here, we present algorithms to estimate the $\overline{e}_k$'s for $N \geq n$ which are based on minimizing a norm of the error. The least squares (LS) solution minimizes the 2-norm of the error. Using linear programming we can either minimize the 1-norm or the ∞-norm of the error. Substituting in (20) distinct $z_k \in \mathbb{R} \setminus 0, k = 1, ..., N$ will lead us to the following set of linear equations of the expected alternating elementary symmetric functions, i.e.,

$$\mathbf{Y} = \mathbf{S}\mathbf{X}, \tag{50}$$

where here $\mathbf{X} := (x_1, ..., x_N)^T := ((-1)^1\overline{e}_1, (-1)^2\overline{e}_2, ..., (-1)^n\overline{e}_n)^T$ and $\mathbf{Y} := (y_1, ..., y_N)^T$. We have

$$y_k = \mathbb{E}(1 - z_k^{-1})^K - 1 \tag{51}$$

and

$$\mathbf{S}(k, \ell) := z_k^{-\ell}, k = 1, .., N, \ell = 1, ..., n. \tag{52}$$

Since we have explicit expressions for $\mathbf{Y}$ we can replace $\mathbf{Y}$ by an estimator, say $\hat{\mathbf{Y}}$, and by using (50) estimate $\mathbf{X}$ either by least squares or by using LP. By using least squares for $N \geq n$ we obtain

$$\hat{\mathbf{X}} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\hat{\mathbf{Y}}, \tag{53}$$

which for $N = n$ reduces to $\hat{\mathbf{X}} = \mathbf{S}^{-1}\hat{\mathbf{Y}}$. Notice that for $N \geq n$, $\mathbf{S}$ must be of full rank because otherwise the polynomial of degree $n-1$ whose coefficients are $\mathbf{X}$ would have more than $n-1$ zeros which is impossible. Notice that $\hat{e}_k = |\hat{\mathbf{X}}(k)|$. Next using LP we can compute $\hat{\mathbf{X}}$ either by solving

$$\min_{\mathbf{X}} ||\hat{\mathbf{Y}} - \mathbf{S}\mathbf{X}||_1, \tag{54}$$

or by solving

$$\min_{\mathbf{X}} ||\hat{\mathbf{Y}} - \mathbf{S}\mathbf{X}||_\infty, \tag{55}$$

where $||(v_1, ..., v_n)||_1 := \sum_{i=1}^n |v_i|$ and $||(v_1, ..., v_n)||_\infty := \max(|v_1|, ..., |v_n|)$).

2.7 The Case When z Is a Random Variable, say Z.

There is no reason why we could not choose z to be a random variable, say Z. We will give two examples for this case. In the first example we choose without loss of generality $Z = \mathbf{1}(A_1)$ and obtain two identities.

In the second example we choose $Z \sim N(0, \sigma^2)$ independent of the A_i's and obtain quite a complex result.

Example 2.6. *Let $Z = \mathbf{1}(A_1)$. Then using (13) and (14) we obtain*

$$(\mathbf{1}(A_1) - 1)^K \mathbf{1}(A_1)^{n-K} \equiv 0, \tag{56}$$

which is otherwise tricky to prove. We also have

$$
\begin{aligned}
0 &\equiv Z^n + \sum_{k=1}^{n-1}(-1)^k \sum_{\alpha\in Q_{k,n}} \mathbf{1}(\cap_{i=1}^k A_{\alpha_i})Z^{n-k} + (-1)^n \mathbf{1}(\cap_{i=1}^n A_i) \quad (57)\\
&= \mathbf{1}(A_1)\left(1 + \sum_{k=1}^{n-1}(-1)^k \sum_{\alpha\in Q_{k,n}} \mathbf{1}(\cap_{i=1}^k A_{\alpha_i})\right) + (-1)^n \mathbf{1}(A_1)\mathbf{1}(\cap_{i=1}^n A_i)\\
&= \mathbf{1}(A_1)\left(1 + \sum_{k=1}^{n-1}(-1)^k \sum_{\alpha\in Q_{k,n}} \mathbf{1}(\cap_{i=1}^k A_{\alpha_i}) + (-1)^n \mathbf{1}(\cap_{i=1}^n A_i)\right).
\end{aligned}
$$

This identity can be written differently as a telescopic series. For $n = 3$ we have:

$$0 \equiv \mathbf{1}(A_1) - [\mathbf{1}(A_1) + \mathbf{1}(A_1\cap A_2) + \mathbf{1}(A_1\cap A_3)] + [\mathbf{1}(A_1\cap A_2) + \mathbf{1}(A_1\cap A_3) + \mathbf{1}(\cap_{i=1}^3 A_i] - \mathbf{1}(\cap_{i=1}^3 A_i)). \quad (58)$$

Example 2.7. *Let $Z \sim N(0,\sigma^2)$ be a zero mean random variable normally distributed with variance σ^2 independent of the events $A_i, i = 1, ..., n$. It is well known that*

$$\mathbb{E}Z^m = \begin{cases} \frac{(2m)!}{2^m m!}\sigma^m, & m \text{ even}\\ 0, & \text{otherwise.}\end{cases} \quad (59)$$

Hence, using (13) and (14) we obtain

$$\mathbb{E}\left[(Z-1)^K Z^{n-K}\right] = \begin{cases} \mathbb{E}Z^n & + \sum_{k\in\{2,4,...,n-2\}}(-1)^k \overline{e}_k \mathbb{E}Z^{n-k} + (-1)^n \overline{e}_n, \quad n \text{ even}\\ 0 & + \sum_{k\in\{1,3,...,n-2\}}(-1)^k \overline{e}_k \mathbb{E}Z^{n-k} + (-1)^n \overline{e}_n, \quad n \text{ odd.}\end{cases} \quad (60)$$

3. Formal Proof of Known Formulas for $\overline{e}_k := \mathbb{E}e_k, k = 0, ..., n$ and Some Consequences

Here, we will provide a formal proof that follows from (20) to a result that appears in (Grimmet & Stirzaker, 2001, p.158, eqn. 13) for which the authors gave a verbal argument of its validity, i.e,.

Theorem 3.1. *Let $q_k := \mathbb{P}(K = k), k = 0, 1, ..., n$. Then*

$$\overline{e}_k = \mathbb{E}\binom{K}{k}, k = 0, 1, ..., n. \quad (61)$$

Proof: Using (30) we obtain

$$
\begin{aligned}
\mathbb{E}(1-y)^K &= \sum_{k=0}^n q_k(1-y)^k \quad (62)\\
&= \sum_{k=0}^n (-1)^k \overline{e}_k y^k.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\sum_{k=0}^n q_k(1-y)^k &= \sum_{k=0}^n q_k\left(\sum_{\ell=0}^k (-1)^\ell \binom{k}{\ell} y^\ell\right) \quad (63)\\
&= \sum_{\ell=0}^n (-1)^\ell \left(\sum_{k=\ell}^n q_k\binom{k}{\ell}\right) y^\ell\\
&= \sum_{\ell=0}^n (-1)^\ell \overline{e}_\ell y^\ell.
\end{aligned}
$$

Equating coefficients of y^ℓ on both sides of the last two equations we obtain

$$\begin{aligned}\overline{e}_\ell &= \sum_{k=\ell}^{n} q_k \binom{k}{\ell} \\ &= \sum_{k=0}^{n} q_k \binom{k}{\ell}, \text{ since } \binom{k}{\ell} := 0, \ell > k \\ &= \mathbb{E}\binom{K}{\ell}, \ell = 0, 1, ..., n.\end{aligned} \tag{64}$$

This completes the proof. ■

The above expressions for $\overline{e}_k$ and q_k can be written in matrix form as follows.

$$\begin{pmatrix}\overline{e}_0\\ \overline{e}_1\\ \vdots\\ \overline{e}_n\end{pmatrix} = \begin{pmatrix}\binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \cdots & \binom{n}{0}\\ 0 & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \cdots & \binom{n}{1}\\ 0 & 0 & \binom{2}{2} & \binom{3}{2} & \cdots & \binom{n}{2}\\ 0 & 0 & 0 & \binom{3}{3} & \cdots & \binom{n}{3}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & \binom{n}{n}\end{pmatrix}\begin{pmatrix}q_0\\ q_1\\ \vdots\\ q_n\end{pmatrix}. \tag{65}$$

Let $\overline{\mathbf{e}}_{n+1} := (\overline{e}_0, \overline{e}_1, ..., \overline{e}_n)^T$, $\mathbf{q}_{n+1} := (q_0, q_1, ..., q_n)^T$, where T denotes transposition, and $\mathbf{P}_{n+1}$ denote the above $(n+1)$x$(n+1)$ upper triangular Pascal matrix, then

$$\overline{\mathbf{e}}_{n+1} = \mathbf{P}_{n+1}\mathbf{q}_{n+1}, \tag{66}$$

where

$$\mathbf{P}_{n+1}(i, j) = \binom{i-1}{j-1}\mathbf{1}(i \le j). \tag{67}$$

Notice that each column of $\mathbf{P}_{n+1}$ can be easily computed by additions of elements of the previous column by using Pacal's triangle formulas. Multiplying both sides of this equation by $\mathbf{e} := (1, 1, ..., 1) \in \mathbb{R}^{n+1}$ we obtain another proof of (29), i.e.,

$$\sum_{k=0}^{n} \overline{e}_k = \mathbf{e}\mathbf{P}_{n+1}\mathbf{q}_{n+1} = \sum_{k=0}^{n} q_k 2^k = \mathbb{E}(2^K). \tag{68}$$

Using $\mathbf{q}_{n+1} = \mathbf{P}_{n+1}^{-1}\overline{\mathbf{e}}_{n+1}$ and Waring's formulas (33) we obtain

$$\mathbf{P}_{n+1}^{-1}(i, j) = (-1)^{i+j}\binom{i-1}{j-1}\mathbf{1}(i \le j). \tag{69}$$

Hence, $\mathbf{P}_{n+1}(i, j)$ is also the (j, i) minor of $\mathbf{P}_{n+1}$.

For example, when $n = 4$ we obtain

$$\mathbf{P}_5^{-1} = \begin{pmatrix}\binom{0}{0} & -\binom{1}{0} & \binom{2}{0} & -\binom{3}{0} & \binom{4}{0}\\ 0 & \binom{1}{1} & -\binom{2}{1} & \binom{3}{1} & -\binom{4}{1}\\ 0 & 0 & \binom{2}{2} & -\binom{3}{2} & \binom{4}{2}\\ 0 & 0 & 0 & \binom{3}{3} & -\binom{4}{3}\\ 0 & 0 & 0 & 0 & \binom{4}{4}\end{pmatrix} = \begin{pmatrix}1 & -1 & 1 & -1 & 1\\ 0 & 1 & -2 & 3 & -4\\ 0 & 0 & 1 & -3 & 6\\ 0 & 0 & 0 & 1 & -4\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}. \tag{70}$$

4. Newtons Identities and Some Consequences

This Section consists of two subsections. In the first subsection we will explore the elementary symmetric functions in terms of the power sums. In the second subsection we will explore the complete homogeneous symmetric functions h_k's in terms of the power sums.

4.1 Newton's Identities Relating the $\overline{e}_k$'s to the Moments $\mathbb{E}(K^k)$ for $k = 1, ..., n$ and Some Consequences

Here, we will explore Newton's identities that relate the power sums to the elementary symmetric functions (Wikipedia contributions, 2021, Feb. 4).

Substituting $\lambda_k = \mathbf{1}(A_k), k = 1, .., n$ in the power sums $p_k, k \geq 1$ we obtain

$$p_k = \sum_{i=1}^{n} (\mathbf{1}(A_i))^k = \sum_{i=1}^{n} \mathbf{1}(A_i) = e_1 \equiv K, k \geq 1. \tag{71}$$

Hence using (Wikipedia contributions, 2021, Feb. 4) we obtain

$$\begin{aligned} e_1 &= K, \\ e_2 &= \frac{1}{2}K^2 - \frac{1}{2}K, \\ e_3 &= \frac{1}{6}K^3 - \frac{1}{2}K^2 + \frac{1}{3}K, \\ e_4 &= \frac{1}{24}K^4 - \frac{1}{4}K^3 + (\frac{1}{8} + \frac{1}{3})K^2 - \frac{1}{4}K, \\ &\vdots \\ e_n &= (-1)^n \sum_{m_1+2m_2+...+nm_n=n} \prod_{i=1}^{n} \frac{(-K)^{m_i}}{m_i! i^{m_i}} \ \forall m_i \geq 0. \end{aligned} \tag{72}$$

Notice that since in the equation for e_4 the coefficient of K^2 is the sum of two values, the equation for e_n can be further simplified.

Taking expectations we obtain

$$\begin{aligned} \overline{e}_1 &= \mathbb{E}(K), \\ \overline{e}_2 &= \frac{1}{2}\mathbb{E}(K^2) - \frac{1}{2}\mathbb{E}(K), \\ \overline{e}_3 &= \frac{1}{6}\mathbb{E}(K^3) - \frac{1}{2}\mathbb{E}(K^2) + \frac{1}{3}\mathbb{E}(K), \\ \overline{e}_4 &= \frac{1}{24}\mathbb{E}(K^4) - \frac{1}{4}\mathbb{E}(K^3) + \frac{11}{24}\mathbb{E}(K^2) - \frac{1}{4}\mathbb{E}(K), \\ &\vdots \\ \overline{e}_n &= (-1)^n \sum_{m_1+2m_2+...+nm_n=n} \prod_{i=1}^{n} \frac{\mathbb{E}[(-K)^{m_i}]}{m_i! i^{m_i}}, \text{all } m_i \geq 0. \end{aligned} \tag{73}$$

Hence, since

$$\overline{e}_k = \overline{e}_k \left(\mathbb{E}(K), \mathbb{E}(K^2), ..., \mathbb{E}(K^k) \right), k = 1, ..., n, \tag{74}$$

we obtain that the $\overline{e}_k$'s are linearly related to the $\mathbb{E}(K^k)$'s via a lower triangular matrix.

In what follows we will show how to obtain the above equations for $\overline{e}_k, k = 1, ..., n$ via matrix multiplication of a submatrix of $\mathbf{P}_{n+1}$ by the inverse of a matrix $\mathbf{G}_n$ whose columns are consecutive geometric series. In particular, we could thus validate the equations for $\overline{e}_1, ..., \overline{e}_4$.

Let $\mathbf{L}_n$ denote the yet unknown lower triangular matrix that relates $\tilde{\mathbf{e}}_n := (\overline{e}_1, ..., \overline{e}_n)^T$ to $\mathbf{m}_n := (\mathbb{E}(K^1), ...\mathbb{E}(K^n))^T$, i.e., $\tilde{\mathbf{e}}_n = \mathbf{L}_n \mathbf{m}_n$. Using $\mathbb{E}(K^\ell) = \sum_{k=0}^{n} k^\ell q_k = \sum_{k=1}^{n} k^\ell q_k$ we obtain

$$\mathbf{m}_n = \mathbf{G}_n \tilde{\mathbf{q}}_n, \tag{75}$$

where $\tilde{\mathbf{q}}_n = (q_1, ., , q_n)$ and

$$\mathbf{G}_n := \begin{pmatrix} 1^1 & 2^1 & \cdots & n^1 \\ 1^2 & 2^2 & \cdots & n^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1^n & 2^n & \cdots & n^n \end{pmatrix}. \tag{76}$$

We thus obtain

$$\mathbf{m}_n = \mathbf{G}_n \tilde{\mathbf{q}}_n. \tag{77}$$

Hence,

$$\tilde{\mathbf{e}}_n = \mathbf{L}_n \mathbf{m}_n = \mathbf{L}_n \mathbf{G}_n \tilde{\mathbf{q}}_n, \tag{78}$$

where we also have $\tilde{\mathbf{e}}_n := \overline{\mathbf{e}}_{n+1}(2, ..., n+1)$ and $\tilde{\mathbf{q}}_n := \mathbf{q}_{n+1}(2, ..., n+1)$. Now, by using (66) we obtain

$$\tilde{\mathbf{e}}_n = \tilde{\mathbf{P}}_n \tilde{\mathbf{q}}_n, \tag{79}$$

where we obtain $\tilde{\mathbf{P}}_n$ from $\mathbf{P}_{n+1}$ by deleting its first row and its first column. Note that since $\mathbf{P}_{n+1}$ is an upper triangular matrix so is its inverse, therefore, we can obtain $\tilde{\mathbf{P}}_n^{-1}$ from $\mathbf{P}_{n+1}^{-1}$ by deleting its first row and its first column. Combining (78) and (79) we obtain $\mathbf{L}_n \mathbf{G}_n = \tilde{\mathbf{P}}_n$. Hence,

$$\mathbf{L}_n = \tilde{\mathbf{P}}_n \mathbf{G}_n^{-1}. \tag{80}$$

Remark 4.1. *It is a rare event that multiplication of matrices as in (80) gives a lower triangular matrix. Hence, we obtain the following $n(n-1)/2$ non trivial identities relating binomial coefficients to elements of $\mathbf{G}_n^{-1}$, i.e.,*

$$[\tilde{\mathbf{P}}_n \mathbf{G}_n^{-1}]_{k,\ell} = 0, \ 1 \le k < \ell \le n. \tag{81}$$

4.2 Newton's Identities Relating the $\overline{h}_k$'s to to the $\overline{e}_k$'s via the Transpose of Pascal's Matrix and Some Consequences

Substituting in (9) $\lambda_\ell = \mathbf{1}(A_\ell), \ell = 1, .., n$ the complete homogeneous symmetric functions become

$$\begin{aligned} h_k(\lambda_1, ..., \lambda_n) &= \sum_{\alpha \in G_{k,n}} \prod_{i=1}^{k} \mathbf{1}(A_{\alpha_i}), \\ &= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \sum_{\alpha \in Q_{\ell,n}} \prod_{i=1}^{k} \mathbf{1}(A_{\alpha_i}) \\ &= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \sum_{\alpha \in Q_{\ell,n}} \mathbf{1}\left(\cap_{i=1}^{\ell} A_{\alpha_i}\right). \end{aligned} \tag{82}$$

The term $\binom{k-1}{\ell-1}$ is the number of monomials of degree k that after the substitution $\lambda_i = \mathbf{1}(A_i), i = 1, ..., n$ become of degree $\ell \le k$. Of the k indices we are left with $k - \ell$ indices that can be distributed among the ℓ indices in $|G_{\ell,k-\ell}| = \binom{k-1}{\ell-1}$ ways.

Remark 4.2. *Notice that if we remove $\binom{k-1}{\ell-1}$ from (82) we obtain another symmetric function.*

Taking expectations we obtain

$$\begin{aligned} \overline{h}_k &= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \sum_{\alpha \in Q_{\ell,n}} \mathbb{E}\mathbf{1}\left(\cap_{i=1}^{\ell} A_{\alpha_i}\right) \\ &= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \sum_{\alpha \in Q_{\ell,n}} \mathbb{P}\left(\cap_{i=1}^{\ell} A_{\alpha_i}\right) \\ &= \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} \overline{e}_\ell. \end{aligned} \tag{83}$$

Let $\overline{\mathbf{h}}_n := (\overline{h}_1, ..., \overline{h}_n)^T$ and $\tilde{\mathbf{e}}_n := (\overline{e}_1, ..., \overline{e}_n)^T$, then the last equation can be written in matrix form as

$$\overline{\mathbf{h}}_n = \mathbf{B}_n \tilde{\mathbf{e}}_n, \tag{84}$$

where

$$\mathbf{B}_n(i,j) = \binom{i-1}{j-1}, i,j = 1,...,n \tag{85}$$

$$= \binom{i-1}{j-1}\mathbf{1}(i \geq j). \tag{86}$$

Notice that $\mathbf{B}_n$ is a lower triangular matrix with a diagonal of ones and its k's row is composed of the binomial coefficients the expansion of $(1+1)^{k-1}$.

For example, when n=5 using (70) we obtain

$$\mathbf{B}_5 := \begin{pmatrix} \binom{1-1}{0} & 0 & 0 & 0 & 0 \\ \binom{2-1}{0} & \binom{2-1}{1} & 0 & 0 & 0 \\ \binom{3-1}{0} & \binom{3-1}{1} & \binom{3-1}{2} & 0 & 0 \\ \binom{4-1}{0} & \binom{4-1}{1} & \binom{4-1}{2} & \binom{4-1}{3} & 0 \\ \binom{5-1}{0} & \binom{5-1}{1} & \binom{5-1}{2} & \binom{5-1}{3} & \binom{5-1}{4} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix} = \mathbf{P}_5^T. \tag{87}$$

and

$$\mathbf{B}_5^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix}. \tag{88}$$

More generally, using (67) we obtain

$$\mathbf{B}_n = \mathbf{P}_n^T. \tag{89}$$

Hence,

$$\mathbf{B}_n^{-1} = (\mathbf{P}_n^T)^{-1} = (\mathbf{P}_n^{-1})^T. \tag{90}$$

So, the elements of the inverse of $\mathbf{B}_n$ are given by

$$\mathbf{B}_n^{-1}(i,j) = (-1)^{i+j}\binom{i-1}{j-1}, i,j = 1,...,n. \tag{91}$$

5. Inequalities Associated with K and Various Symmetric Functions

In what follows we will present four examples of pertinent inequalities. In the last example we will state Kounias's and Bonferroni's inequalities (Grimmet & Stirzaker, 2001, p. 25). We will propose a conjecture associated Bonferroni's inequality that also holds for all $n > 5$. If this conjecture is true it will give a series of alternating higher order Bonferroni type inequalities that in turn may give a series of alternating higher order Kounias type inequalities.

Example 5.1. *Power means (Marcus & Minc, 1964, p.105):*

$$\mathcal{M}_r := \begin{cases} \left(\prod_{i=1}^n \lambda_i\right)^{1/n}, & r = 0 \\ \left(\frac{\sum_{i=1}^n \lambda_i^r}{n}\right)^{1/r}, & r \neq 0. \end{cases} \tag{92}$$

For fixed $(\lambda_1, ..., \lambda_n)$ *and* $r < s$ *we have (Marcus & Minc, 1964, p.105):*

$$\mathcal{M}_r \leq \mathcal{M}_s. \tag{93}$$

Let $\lambda_i = \mathbf{1}(A_i), i = 1,...,n$*. Hence, for* $s > r = 0$ *we obtain*

$${e_n}^{1/n} \leq \left(\frac{K}{n}\right)^{1/s}. \tag{94}$$

After some algebra and taking expectations we obtain

$$\overline{e}_n \leq \frac{\mathbb{E}K^{n/s}}{n^{n/s}}, s > 0. \tag{95}$$

Notice that the case $s > r > 0$ gives a trivial result.

Example 5.2. *Let (Marcus & Minc, 1964, p.106)*

$$p_k := e_k \binom{n}{k}^{-1} \tag{96}$$

denote the k^{th} weighted elementary symmetric function of the $\mathbf{1}(A_i)$'s. Then (Marcus & Minc, 1964, p.106):

$$p_1 > p_2^{1/2} > p_3^{1/3} > \cdots > p_n^{1/n}. \tag{97}$$

Since,

$$p_1 > p_k^{1/k}, k > 1 \tag{98}$$

we obtain

$$\left(\frac{K}{n}\right)^k > e_k \binom{n}{k}^{-1}, \tag{99}$$

which after some algebra and taking expectations we obtain

$$\overline{e}_k < n^{-k}\binom{n}{k}\mathbb{E}K^k. \tag{100}$$

Example 5.3. *For $(a_1, ..., a_n)$ and $(b_1, ..., b_n)$ vectors in the nonnegative orthant of $\mathbb{R}^n$ and $0 < \theta < 1$, Minkowsky's inequality states (Marcus and Minc, 1964, p.109):*

$$\left(\sum_{i=1}^n (a_i + b_i)^{1/\theta}\right)^\theta \leq \left(\sum_{i=1}^n a_i^{1/\theta}\right)^\theta + \left(\sum_{i=1}^n b_i^{1/\theta}\right)^\theta. \tag{101}$$

Substituting $a_i = \mathbf{1}(A_i)$ and $b_i := \mathbf{1}(A_i^c), i = 1, ..., n$ and noting that $\mathbf{1}(A_i) + \mathbf{1}(A_i^c) = 1, i = 1, ..., n$, we obtain

$$\begin{aligned} n^\theta \leq& \left(\sum_{i=1}^n \mathbf{1}(A_i)\right)^\theta + \left(\sum_{i=1}^n \mathbf{1}(A_i^c)\right)^\theta \\ =& K^\theta + (n-K)^\theta, K \in \{0, 1, ..., n\}, \theta \in (0, 1]. \end{aligned} \tag{102}$$

Taking expectations we obtain

$$\mathbb{E}(K^\theta) + \mathbb{E}((n-K)^\theta) \geq n^\theta, \theta \in (0, 1]. \tag{103}$$

Example 5.4. *For the sake of completeness we mention Kounias's and Bonferroni's inequalities [1, p. 25]. Their proofs by induction appears in (Grimmet & Stirzaker, 2003, p. 150). Then, based on the nonnegativity of probability, the union bound, and Bonferroni's inequality we will propose a conjecture.*

(i) Kounias's inequality

$$\mathbb{P}(\cup_{i=1}^n A_i) \leq \min_k \left(\sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i:i\neq k} \mathbb{P}(A_i \cap A_k)\right). \tag{104}$$

(ii) Bonferroni's inequality

$$\begin{aligned} \mathbb{P}(\cup_{i=1}^n A_i) \geq& \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{\alpha \in Q_{2,n}} \mathbb{P}(A_{\alpha_1} \cap A_{\alpha_2}). \\ =& \overline{e}_1 - \overline{e}_2. \end{aligned} \tag{105}$$

We will now show how to obtain the proposed conjecture.

(1) Since $\mathbb{P}(\cup_{i=1}^{n} A_i) \geq 0$, *using (26) we obtain*

$$\overline{e}_1 \geq \overline{e}_2 - \overline{e}_3 + ... + (-1)^{n-2}\overline{e}_n. \tag{106}$$

(2) Using the union bound and (26) we obtain

$$\overline{e}_2 \geq \overline{e}_3 - \overline{e}_4 + ... + (-1)^{n-3}\overline{e}_n. \tag{107}$$

(3) Using Bonferroni's inequality and (26) we obtain

$$\overline{e}_3 \geq \overline{e}_4 - \overline{e}_5 + ... + (-1)^{n-4}\overline{e}_n. \tag{108}$$

Hence, we conjecture that this pattern continues until

$$\overline{e}_{n-1} \geq \overline{e}_n \tag{109}$$

Notice that the last inequality holds since it obviously holds for each addend of $\overline{e}_{n-1}$. *Since we proved the prposed conjecture for* $n \leq 5$ *it needs to be resolved only for* $n > 5$. *For example, if the proposed conjecture holds we obtain for* $n > 5$

$$\mathbb{P}(\cup_{i=1}^{n} A_i) \leq \overline{e}_1 - \overline{e}_2 + \overline{e}_3. \tag{110}$$

Area for further research would be to use the proposed methodology to find new proofs to the above two inequalities and to resolve the proposed conjecture.

6. Exact Polynomial Time Algorithm for Computing the Expected Elementary Symmetric Functions for Finite Sample Spaces

The complexity of computing all the $\overline{e}_k$'s is $O(2^n)$ since we need to compute all the $2^n - 1$ probabilities of the intersection of all the subsets of $\{A_1, ..., A_n\}$, excluding the empty set. Notice that n can reach values of $O(2^{|\Omega|})$ and then the complexity becomes $O(2^{2^{|\Omega|}})$). The proposed algorithm is based on (65), (66) where the vector of $\overline{e}_k$'s, $\overline{\mathbf{e}}_{n+1}$, is obtained by multiplying Pascal's upper triangular matrix, $\mathbf{P}_{n+1}$ by the vector of the q_k's, $\mathbf{q}_{n+1}$. We can compute and store a matrix $\mathbf{P}_N$ for N large enough and then retrieve from memory the leading submatrix $\mathbf{P}_{n+1}$ of $\mathbf{P}_N$. Hence, given $\mathbf{q}_{n+1}$, the complexity of computing $\overline{\mathbf{e}}_{n+1}$ is $O(n^2)$. In what follows we will present an algorithm to compute $\mathbf{q}_{n+1}$ whose complexity is $O(n|\Omega|)$. Therefore, the complexity of the proposed algorithm is $O(n(\max\{n, |\Omega|\})$.

Let $\mathbf{A} =\in \{\mathbf{0}, \mathbf{1}\}^{\mathbf{nx|\Omega|}}$ denote the matrix whose elements are

$$a_{i,j} := \begin{cases} 1, & \text{if } \omega_j \in A_i \\ 0, & \text{otherwise.} \end{cases} \tag{111}$$

Hence,

$$A_i = \{\omega_j \in \Omega : a_{i,j} = 1, j = 1, \ldots, |\Omega|\}. \tag{112}$$

Let

$$\mathbf{K} = \text{sum}(\mathbf{A}) := \left(\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{a_{i,1}}, ..., \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{a_{i,|\Omega|}} \right). \tag{113}$$

Let $\pi_i := \mathbb{P}(\omega_i), i = 1, \ldots, |\Omega|$, where $\sum_{i=1}^{|\Omega|} \pi_i = 1$. Notice that if all the ω_i's are equally likely then $\pi_i = 1/|\Omega|$. Hence, we obtain

$$\mathbb{P}(K = k) = \sum_{i=1}^{|\Omega|} \pi_i \mathbf{1}(\mathbf{K}(i) = k). \tag{114}$$

We have

$$\begin{aligned}
&\mathbf{q}_{n+1} = 0\\
&\mathbf{K} = \mathrm{sum}(\mathbf{A})\\
&\text{for } i = 1, ..., |\Omega|\\
&\quad \mathbf{q}(\mathbf{K}(i) + 1) = \mathbf{q}(\mathbf{K}(i) + 1) + \pi_i\\
&\text{end.}
\end{aligned} \tag{115}$$

Hence, the total complexity of computing $\overline{\mathbf{e}}_{n+1}$ is the sum of the complexities of computing $\mathbf{K}$, $\mathbf{q}_{n+1}$, and then $\overline{\mathbf{e}}_{n+1}$, i.e.,

$$O(|\Omega|n) + O(|\Omega|) + O(n^2) = O(n \max\{n, |\Omega|\}). \tag{116}$$

Using $\overline{\mathbf{e}}_{n+1}$ we can compute $\mathbb{P}(\cup_{i=1}^{n} A_i)$ and $\mathbb{E}(2^K)$ with added complexity $O(n)$ which does not alter the above complexity result.

7. Conclusion

The main results of this article are based on (i) taking expectations of the various symmetric functions (elementary, power, and complete symmetric functions) after replacing their independent variables by the the indicator functions $\mathbf{1}(A_i), i = 1, ..., n$ of the events $A_i, i = 1, ..., n$; and, (ii) by using the rather simple observation that $f_n(z) = \prod_{i=1}^{n}(z - \mathbf{1}(A_n)) \equiv (z-1)^K z^{n-K}$, where $z \in \mathbb{C}$ is a parameter and K is the number of A_i's that occured; and, then expanding the LHS by using Vieta's formulas and taking expectations. We gave the following examples. (i) When $z = 0$ we obtained the obvious identity $\mathbb{P}(K = n) \equiv \mathbb{P}(\cap_{1=1}^{n} A_i)$ implying the identity $\mathbb{P}(K = 0) \equiv 1 - \mathbb{P}(\cup_{1=1}^{n} A_i)$. (ii) When $z = 1$ we obtained the inclusion exclusion formula for $\mathbb{P}(\cup_{1=1}^{n} A_i)$ as the sum of alternating expected elementary symmetric functions. (iii) When $z = -1$ we obtained a formula for $\mathbb{E}(2^K)$ as one plus the sum of all the expected elementary symmetric functions. Verbally, $\mathbb{E}(2^K)$ is the expected number of subfamilies of $\{A_1, ..., A_n\}$ all whose component events occur. (iv) When $z^{-1} = 1 - s$, $z^{-1} = 1 - \exp(s)$, and $z^{-1} = 1 - \exp(it)$ we obtained the probability generating, the moment generating, and the characteristic functions of K, respectively. Using the probability generating function thus obtained we gave a simpler proof of Waring's formulas that relates the $q_k := \mathbb{P}(K = k)$'s to the expected elementary symmetric functions, the $\overline{e}_k$'s. (v) When $z_k = \exp(\frac{2\pi i}{N}k), k = 0, 1, ..., N-1, N \geq n, i = \sqrt{-1}$, we presented a least squares efficient algorithm based on an N-point **IFFT** (Inverse Fast Fourier Transform) to estimate the $\overline{e}_k$'s. (vi) When $z_k \in \mathbb{R} \setminus 0, k = 1, ..., N, N \geq n$ are distinct we gave algorithms based on least squares (LS) and linear programming (LP) to estimate the $\overline{e}_k$'s. (vii) When $Z = \mathbf{1}(A_i)$ and $Z \sim N(0, 1)$ is standard normal we arrived at several new identities.
Using the above equation $f_n(z) = (z-1)^K z^{n-K}$, we presented formal proofs of known formulas for the $\overline{e}_k$'s as a function of the q_k's and gave some consequences. Then, we replaced in Newton's identities the independent variables $\lambda_1, ..., \lambda_n$ by $\lambda_k = \mathbf{1}(A_k), k = 1, .., n$, took expectations and gave some consequences. We also replaced in some well known inequalities the independent variables $\lambda_1, ..., \lambda_n$ by $\lambda_k = \mathbf{1}(A_k)$, took expectations and gave some consequences. Moreover, based on the non-negativity of probability, the union bound, and Bonferroni's inequality we showed that $\overline{e}_k \geq \overline{e}_{k-1} - \overline{e}_{k-2} + \cdots + (-1)^{n-k+1}\overline{e}_n$ for $k = 1, 2, 3, n-1$ and conjectured that it is also true for all $3 < k < n-1$ and $n > 5$. Next, we presented a polynomial time algorithm to compute the exact q_k's and $\overline{e}_k$'s for finite sample spaces.

Further research will focus on proving other existing results or obtaining other new results. In particular, finding new proofs of Bonferroni's and Kounias's inequalities by using the above methodology and resolving the conjecture associated with the union bound and Bonferroni's inequality, for more details see Example 5.4. Error analysis of the proposed estimation algorithms is also needed.
Finally, I would like to mention that the above methodology turned out to be one of my great teachers, hopefully of the readers too.

Acknowledgments

The author is grateful to the reviewer for his useful remarks and to Prof. David Malah for suggesting to swap the abstract with the conclusion section.

References

Grimmet, G., & Stirzaker, D. (2001). *Probability and Random Processes* (3rd ed). Oxford University Press.

Grimmet, G., & Stirzaker, D. (2003). *One Thousand Exercises in Probability*. Oxford University Press.

Marcus, M., & Minc, H. (1964). *Survey of Matrix Theory and Matrix Inequalities*. Dover publications.

Wikipedia contributors. (2019, October 19). Symmetric polynomial. In Wikipedia, The Free Encyclopedia. Retrieved 16:23, March 3,2021, from https://en.wikipedia.org/w/index.php?title=Symmetric_polynomial&oldid=922077796

Wikipedia contributors. (2021, February 4). Newton's identities. In Wikipedia, The Free Encyclopedia. Retrieved 16:43, March 3, 2021, from https://en.wikipedia.org/w/index.php?title=Newton%27s_identities&oldid=1004851793

Copyrights