Mahonian Numbers

For what values of $n \geq 1$ does there exist a number m that can be written in the form $a_1 + \ldots + a_n$ (with $a_1 \in \{1\}, a_2 \in \{1, 2\}, \ldots, a_n \in \{1, 2, \ldots, n\}$) in $(n-1)!$ or more ways?

Let us define a function f to return the frequency of a sum k formed by $a_1 + \ldots + a_n$ as

$$f(n, k) = \begin{cases} 0, & \text{if } k < n \\ 1, & \text{if } k = n \\ f\left(n, \dfrac{n(n+3)}{2} - k\right), & \text{if } k > \dfrac{n(n+3)}{4} \\ \sum\limits_{i=1}^{n} f(n-1, k-i) \end{cases}$$

Explanation of cases:

1. $f(n, k) = 0$ if $k < n$: It is impossible to make a selection $(a_1, \ldots, a_n)$ that sums to less than n if $a_i \geq 1$.

2. $f(n, k) = 1$ if $k = n$: There is only 1 selection $(a_1, \ldots, a_n)$ that produces n, which is when all $a_i = 1$.

3. $f(n, k) = f\left(n, \dfrac{n(n+3)}{2} - k\right)$: We can show that for every valid selection $(a_1, a_2, \ldots, a_n)$ that produces k, there is a corresponding selection $(a_1', a_2', \ldots, a_n')$ that produces $\dfrac{n(n+3)}{2} - k$, where

$$a_i' = i + 1 - a_i$$

Note that the mapping of $a_i' = i + 1 - a_i$ is a bijection because

(a) $a_i \in \{1, 2, \ldots, i\}$ and $a_i' \in \{1, 2, \ldots, i\}$

(b) The mapping is invertible as applying the mapping twice over results in

$$a_i'' = i + 1 - a_i' = i + 1 - (i + 1 - a_i) = a_i$$

Looking at the sums $k = a_1 + \ldots + a_n$ and $k' = a'_1 + \ldots + a'_n$, we have

$$k' = \sum_{i=1}^{n} (i + 1 - a_i) = n + \sum_{i-1}^{n} (i - a_i) = n + \frac{n(n+1)}{2} - \sum_{i=1}^{n} a_i = \frac{n(n+3)}{2} - k$$

As we have shown the mapping of $(a_1, \ldots, a_n) \mapsto (a'_1, \ldots, a'_n)$ is bijective, every selection that produces a sum k has a corresponding selection producing the sum $k' = \frac{n(n+3)}{2} - k$. Thus, the frequencies of sums k and $\frac{n(n+3)}{2} - k$ are equal and

$$f(n, k) = f\left(n, \frac{n(n+3)}{2} - k\right)$$

4. $f(n, k) = \sum_{i=1}^{n} f(n - 1, k - i)$: This recurrence relation returns all the ways we can make a target sum k from a selection $(a_1, \ldots, a_n)$ by counting all the ways we can make a target sum $k - i$ from a selection $(a_1, \ldots, a_{n-1})$ for $1 \leq i \leq n$, since we can exclude at least $a_n = 1$ (lower bound of i) and at most $i = a_n = n$ (upper bound of i) from the original selection.

Proof by induction. We can now use induction on n to show that for $n \geq 5$, $f(n, k) < (n - 1)!$.

Base case: $n = 5$. The maximum value of $f(5, k)$ for all k is $22 < (5 - 1)! = 24$. Thus, the base case holds.
Inductive hypothesis: Suppose for all $m \geq 5$, $f(m, k) < (m - 1)!$. Then,

$$\begin{aligned} f(m + 1, k) &= \sum_{i=1}^{m+1} f(m, k - i) \\ &= \sum_{i=1}^{m+1} \sum_{j=1}^{m} f(m - 1, k - i - j) \\ &= \sum_{j=1}^{m} \sum_{i=1}^{m+1} f(m - 1, k - i - j) \end{aligned}$$

For fixed j, the inner sum has $\frac{(m-1)(m-2)}{2}+1$ non-zero terms since $k-i-j$ must be in the range $[m-1, \frac{m(m-1)}{2}]$. So, the number of values of i that contribute a non-zero term to the inner summation is at most

$$\frac{(m-1)(m-2)}{2}+1$$

Then, we can bound $f(m+1,k)$ by

$$\begin{aligned}
f(m+1,k) &\leq \sum_{j=1}^{m}\left(\frac{(m-1)(m-2)}{2}+1\right)\cdot f(m-1,k-j) \\
&\leq \left(\frac{(m-1)(m-2)}{2}+1\right)\cdot \sum_{j=1}^{m} f(m-1,k-j) \\
&\leq \left(\frac{(m-1)(m-2)}{2}+1\right)\cdot (m-1)!
\end{aligned}$$

since the total number of arrangements of length $m-1$ is counted by i choices for a_i, resulting in $(m-1)!$.

Since $\frac{(m-1)(m-2)}{2}+1 < m$ for $m \geq 5$, we can simplify further to

$$f(m+1,k) < m\cdot(m-1)! = m!$$

Thus, the inductive step holds, showing that for all $n \geq 5$, the number of sums that can be written in the form $a_1+\ldots+a_n$ with $a_i \in \{1,\ldots,i\}$ is less than $(n-1)!$. We can quickly verify the cases $n=1$ through $n=4$ as having sum-frequencies of $1=0!, 1=1!, 2=2!$, and $6=3!$ respectively.

□

I did some further analysis into this problem and found that

$$f(n,k) = [x^k]G_n(x)$$

which denotes the coefficient of x^k in the polynomial expansion of the generating function

$$G_n(x) = \prod_{i=1}^{n}\sum_{j=1}^{i} x^j$$

It can be shown that the construction of $G_n(x)$ matches the reasoning provided above for constructing $f(n,k)$.

Furthermore, I found the approximation

$$f(n,k) \approx \sqrt{3}\left(1-\frac{1}{n}\right) \cdot \sum_{m=0}^{\lfloor x \rfloor} (-1)^m \binom{n}{m} (x-m)^{n-1}$$

where

$$x = \operatorname{sgn}(c(n,k)) \cdot |c(n,k)|^{\frac{1}{\log(n) \cdot n}} + \frac{n}{2}$$

and

$$c(n,k) = \frac{k-\mu}{\left(1+\frac{1}{\log(n)}\right) \cdot \frac{n-1}{2}}$$

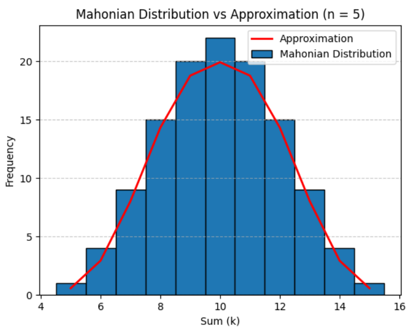
with

$$\mu = \frac{n(n+3)}{4}$$

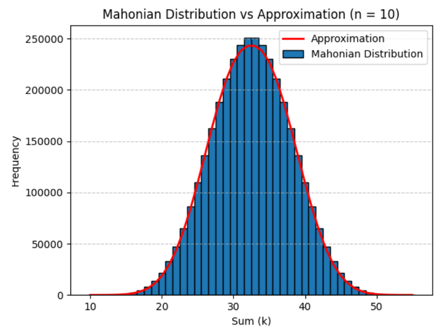
by manipulating the Irwin-Hall distribution based on how f and G_n are constructed, as well as including some constants and functions for continuity correction based on the expected end-behavior of the distribution. Below is a table of the average error of the approximation across all k when compared to the actual values of $f(n,k)$, also known as Mahonian numbers, as well as some graphs showing the approximation superimposed on the distribution.

n	Average error for k
5	9.460981%
10	2.900464%
20	1.177748%
50	0.3454274%
80	0.2891027%

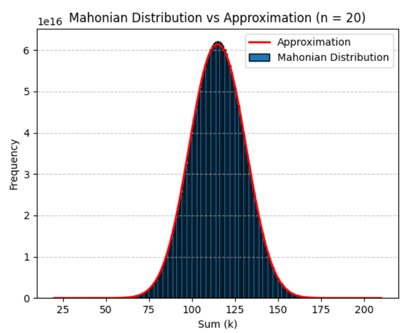
Due to hardware constraints, I was unable to compute error rates or generate plots for larger values of n.



Mahonian approximation for $n = 5$



Mahonian approximation for $n = 10$



Mahonian approximation for $n = 20$