Difference sequences, sums of powers, and Stirling numbers

Difference sequences

Notation:
If $h_0, h_1, ...$ is a number sequence, we will sometimes refer to the sequence just as h.

Definition:
Δ is the operator on number sequences of taking successive differences; for a number sequence h, the number sequence Δh is defined by

$$\Delta h_n = h_{n+1} - h_n.$$

$\Delta^2 h_n = \Delta\Delta h_n$, etc.

We write out a sequence and its iterated differences as an infinite triangle; e.g. if $h_n = n^2$, the iterated differences are as follows:

```
1   4   9   16  25 ...
  3   5   7   9 ...
    2   2   2 ...
      0   0 ...
        0 ...
```

Remark:
Δ is a linear operator, i.e. for sequences h and h' and numbers c and c',

$$\Delta(ch + c'h')_n = c\Delta h_n + c'\Delta h'_n.$$

Hence the powers Δ^k are also linear.

Lemma:
Let f be a polynomial of degree at most d, and let $h_n = f(n)$ be the sequence of its values on natural numbers.

Then $\Delta^{d+1} h_n = 0$ for all n.

Proof:
By linearity, it suffices to show this for monomials $f(x) = x^d$.

So let $h_n = n^d$, and suppose inductively that the lemma holds for polynomials of degree less than d.

Then

$$\Delta h_n = h_{n+1} - h_n = (n+1)^d - n^d$$
$$= n^d + dn^{d-1} + \binom{d}{2} n^{d-2} + ... + 1 - n^d$$

$= \binom{d}{1}n^{d-1} + \binom{d}{2}n^{d-2} + ... + 1$
which has degree $d-1$.

So by the inductive hypothesis,
$0 = \Delta^d \Delta h_n = \Delta^{d+1} h_n$. □

Now suppose $h_n = f(n)$ with f a polynomial of degree d.
By the above lemma, the numbers $h_0, \Delta h_0, ..., \Delta^d h_0$ determine the whole sequence h,
since we can generate the whole triangle from the initial diagonal $h_0, \Delta h_0, ..., \Delta^d h_0, 0, 0, ...$.

Let's find a formula for h_n in terms of $h_0, \Delta h_0, ..., \Delta^d h_0$.

Generating the triangle is a linear process,
so if we can find a formula for h_n generated from an initial diagonal $0, 0, ..., 0, 1, 0, 0, ...$,
with $\Delta^k h_0 = 1$ and all other $\Delta^i h_0 = 0$,
we can then take a linear combination.

We get a "twisted Pascal's triangle", e.g.:

```
0   0   0   0   1   5   15  35  70 ...
  0   0   0   1   4   10  20  35 ...
    0   0   1   3   6   10  15 ...
      0   1   2   3   4   5 ...
        1   1   1   1   1 ...
          0   0   0   0 ...
```

and so we see that $h_n = \binom{n}{k}$.

To prove this: let $f(x) := \frac{x(x-1)(x-2)...(x-(k-1))}{k!}$;
then $f(0) = f(1) = ... = f(k-1) = 0$ and $f(k) = 1$,
so the difference triangle of $f(n)$ also starts with

```
0   0   0   0   1
  0   0   0   1
    0   0   1
      0   1
        1
```

, and since by the lemma it also has 0s thereafter,
we must have $h_n = f(n)$.

Then we directly calculate that $f(n) = \binom{n}{k}$.

So, taking linear combinations, we conclude :

Theorem:
If the initial diagonal of the difference triangle of h_n is $c_0, c_1, ..., c_d, 0, 0, ...$
(i.e. if $\Delta^k h_0 = c_i$ for $k \leq d$, and $\Delta^k h_0 = 0$ for $k > d$),
then
$h_n = \sum_{k=0}^{d} c_k \binom{n}{k}$.

Sums of powers

We can use this theorem to give neat formulae for sums of powers $\sum_{n=0}^{k} n^d$, generalising the formulae you know and love for $d = 1$ and $d = 2$ (and maybe even $d = 3$, if you're that generous with your affections),
and more generally to give formulae for $\sum_{n=0}^{k} f(n)$ where f is any polynomial.

First recall the formula (from the section on Binomial Coefficients)
$\binom{k+1}{r+1} = \sum_{n=0}^{k} \binom{n}{r}$.

So to find $\sum_{n=0}^{k} f(n)$, we can first use the above theorem to find an expression for $f(n)$ in terms of binomial coefficients, then use this formula to sum them.

Example:
Let's find a formula for $\sum_{n=0}^{k} n^4$.

Drawing the start of the difference triangle,

```
 0    1     16   81   256
   1     15   65   175
      14   50   110
         36   60
            24
```

, and recalling that all further rows are 0 since n^4 has degree 4,
we see that the initial diagonal is $0, 1, 14, 36, 24, 0, 0, \ldots$.

So by the above theorem,
$n^4 = \binom{n}{1} + 14\binom{n}{2} + 36\binom{n}{3} + 24\binom{n}{4}$.

So using the formula
$\binom{k+1}{r+1} = \sum_{n=0}^{k} \binom{n}{r}$,
we find
$\sum_{n=0}^{k} n^4$
$= \sum_{n=0}^{k} \left(\binom{n}{1} + 14\binom{n}{2} + 36\binom{n}{3} + 24\binom{n}{4}\right)$
$= \binom{k+1}{2} + 14\binom{k+1}{3} + 36\binom{k+1}{4} + 24\binom{k+1}{5}$

Exercise:
Repeat this procedure for n^1, n^2 and n^3, and check that the answers you get agree with the standard formulae.

Stirling numbers

We would like to understand the mysterious numbers which appear in the formula for $\sum_{n=0}^{k} n^p$, i.e the numbers $c(p,k)$ defined by

$c(p,k) := \Delta^k h_0$ where $h_n = n^p$.

So as we saw, these are the numbers $c(p,k)$ such that

$n^p = \sum_{k=0}^{p} c(p,k)\binom{n}{k}$.

We observe (and will eventually prove) that $c(p,k)$ seems to be divisible by $k!$, so set

$S(p,k) := c(p,k)/k!$.

So, introducing the notation $[n]_k := P(n,k) = k!\binom{n}{k}$,

$n^p = \sum_{k=0}^{p} S(p,k)[n]_k$.

These numbers $S(p,k)$ are the Stirling numbers of the second kind.

Here's a table, written in Pascal triangle format with k going across and p going down, and starting with $S(1,1) = 1$:

```
          1
        1   1
      1   3   1
    1   7   6   1
  1   15  25  10  1
1   31  90  65  15  1
```

This corresponds to the formulae

$n^1 = [n]_1$
$n^2 = [n]_1 + [n]_2$
$n^3 = [n]_1 + 3[n]_2 + [n]_1$
...

All values of $S(p,k)$ not shown in the triangle are 0, except $S(0,0) = 1$.

Lemma:
For all $p > 0$, and all k,

$S(p,k) = S(p-1,k-1) + kS(p-1,k)$.

Proof:
First, note that $S(p,k) = 0$ when $k > p$, by considering degrees of polynomials.
Also $S(p,k) = 0$ when $k < 0$, by definition.

Now

$$
\begin{aligned}
n^p = nn^{p-1} &= n\sum_{k=0}^{p-1} S(p-1,k)[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)((n-k)+k)[n]_k \\
&= \sum_{k=0}^{p-1} S(p-1,k)[n]_{k+1} + \sum_{k=0}^{p-1} kS(p-1,k)[n]_k \\
&= \sum_{k=1}^{p} S(p-1,k-1)[n]_k + \sum_{k=0}^{p-1} kS(p-1,k)[n]_k \\
&= \sum_{k=0}^{p} S(p-1,k-1)[n]_k + \sum_{k=0}^{p} kS(p-1,k)[n]_k \\
&\qquad \textit{(using } S(p-1,-1) = 0 = S(p-1,p)\textit{)} \\
&= \sum_{k=0}^{p}(S(p-1,k-1) + kS(p-1,k))[n]_k,
\end{aligned}
$$

so we conclude by comparing coefficients with
$n^p = \sum_{k=0}^{p} S(p,k)[n]_k$. □

Theorem:
$S(p,k)$ is the number of partitions of a set of p objects into k indistinguishable boxes in which no box is empty,
i.e. the number of partitions of a set of size p into a set of k non-empty subsets,
i.e. the number of sets of non-empty subsets of $\{1,...,p\}$ which are disjoint and have union $\{1,...,p\}$.

Proof:
Write $S'(p,k)$ for this number.

Suppose $p \geq 1$ and $1 \leq k \leq p$.
Consider a partition of $\{1,...,p\}$ into a set of k non-empty subsets,
and consider removing p.
First, suppose the set in the partition which contains p is just $\{p\}$.
Then on removing p, we obtain a partition of $\{1,...,p-1\}$ into $k-1$ subsets.
Otherwise, on removing p we obtain a partition of $\{1,...,p-1\}$ into k subsets.
In the first case, the map is bijective, but in the second case there are k ways of obtaining the same partition of $\{1,...,p-1\}$, since p could have been removed from any of the k sets in that partition.

So
$S'(p,k) = S'(p-1,k-1) + kS'(p-1,k)$.
Clearly $S'(p,k) = 0$ for $k < 0$ or $k > p$ or $p < 0$, and $S(0,0) = 1$.
So by induction on p, $S(p,k) = S'(p,k)$ for all p and k.
□

So now we know that $S(p,k)$ is an integer.

Moreover, we can now reason combinatorially to find a formula for $S(p,k)$:

Theorem:
For $p \geq 0$ and $0 \leq k \leq p$,

$$S(p,k) = \sum_{i=0}^{k} (-1)^i \frac{(k-i)^p}{i!(k-i)!}$$

Proof:
Fix p and k.

Let P be the number of partitions of $\{1,...,p\}$ into an **ordered sequence** of k non-empty subsets.
So $P = k!S(p,k)$.

A partition of $\{1,...,p\}$ into an ordered sequence of k subsets,
with no restrictions on the subsets being non-empty,
just corresponds to a k-colouring of $\{1,...,p\}$,

i.e. a choice of which of the k sets in the partition each element should go in,
so there are k^p such partitions.

Let A_i be the partitions of $\{1, ..., p\}$ into an ordered sequence of k subsets, where the ith is empty.

Such a partition corresponds to a partition into $k-1$ possibly empty subsets, by ignoring the one which is required to be empty.
So $|A_i| = (k-1)^p$.

Similarly, $|A_i \cap A_j| = (k-2)^p$ for $i \neq j$, and generally $|\bigcap_{i \in I} A_i| = (k - |I|)^p$.

So by inclusion-exclusion,

$$\begin{aligned} S(p, k) &= \tfrac{1}{k!} P \\ &= \tfrac{1}{k!}(k^p - |\textstyle\bigcup_i A_i|) \\ &= \tfrac{1}{k!}(k^p - \textstyle\sum_{\emptyset \neq I \subseteq \{1,...,k\}} (-1)^{|I|-1} |\bigcap_{i \in I} A_i|) \\ &= \tfrac{1}{k!}(k^p - \textstyle\sum_{i=1}^k (-1)^{i-1} \binom{k}{i} (k-i)^p) \\ &= \tfrac{1}{k!}(\textstyle\sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^p) \\ &= \textstyle\sum_{i=0}^k (-1)^i \frac{(k-i)^p}{i!(k-i)!}) \end{aligned}$$

□