The text mentions Stirling numbers briefly but does not go into them in any depth. However, they are fascinating numbers with a lot of interesting properties, so I thought I would post a handout about them. This is just for fun and mainly for those who may be interested. You are not required to know the material of this handout, except you should at least know the definition of the Stirling numbers of the second kind and how they are used in counting. That material is in your text.

## Stirling Numbers of the First Kind

The *falling factorial polynomial* of degree $n$ is

$$(x)_n = x(x-1)(x-2)(x-3)\cdots(x-n+1),$$

a polynomial of degree $n$ in one indeterminate $x$. If we evaluate the polynomial at $m$, we get the number of $n$-permutations chosen from a set of size $m$:

$$(m)_n = m(m-1)(m-2)(m-3)\cdots(m-n+1) = \frac{m!}{(m-n)!} = P(m,n).$$

Here are the first few of these polynomials.

$$\begin{aligned}
(x)_0 &= 1 \text{ (the empty product!)}\\
(x)_1 &= x\\
(x)_2 &= x(x-1) = x^2 - x\\
(x)_3 &= x(x-1)(x-2) = x^3 - 3x^2 + 2x\\
(x)_4 &= x(x-1)(x-2)(x-3) = x^4 - 6x^3 + 11x^2 - 6x\\
(x)_5 &= x(x-1)(x-2)(x-3)(x-4) = x^5 - 10x^4 + 35x^3 - 50x^2 + 24x
\end{aligned}$$

The coefficients appearing in $(x)_n$ are called *Stirling numbers of the first kind.* The coefficient of $x^k$ in $(x)_n$ is denoted $s(n,k)$, thus

$$(x)_n = \sum_{k=0}^{n} s(n,k)x^k.$$

The absolute value of $s(n,k)$ is denoted $|s(n,k)|$ and is called an *unsigned Stirling number of the first kind.* The signs alternate, so $s(n,k) = (-1)^{n-k}|s(n,k)|$.

The signed and unsigned Stirling numbers of the first kind satisfy Pascal-like recurrence relations:

(i) $s(n,n) = 1$ for all $n \geq 0$

(ii) $s(n,0) = 0$ for all $n \geq 1$

(iii) $s(n,k) = s(n-1,k-1) - (n-1)\cdot s(n-1,k)$ for $0 < k < n$

(i′) $|s(n,n)| = 1$ for all $n \geq 0$

(ii′) $|s(n,0)| = 0$ for all $n \geq 1$

(iii′) $|s(n,k)| = |s(n-1,k-1)| + (n-1)\cdot|s(n-1,k)|$ for $0 < k < n$.

Arranging the numbers in Pascal-like triangles, the recurrences (iii) and (iii′) say how to obtain an interior entry from the two entries immediately above it. For example, $35 = 11 + 4 \cdot 6$.

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 0 & & 1 & & & & & \\
 & & & & 0 & & -1 & & 1 & & & & \\
 & & & 0 & & 2 & & -3 & & 1 & & & \\
 & & 0 & & -6 & & 11 & & -6 & & 1 & & \\
 & 0 & & 24 & & -50 & & 35 & & -10 & & 1 & \\
0 & & -120 & & 274 & & -225 & & 85 & & -15 & & 1
\end{array}
\qquad
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 0 & & 1 & & & & & \\
 & & & & 0 & & 1 & & 1 & & & & \\
 & & & 0 & & 2 & & 3 & & 1 & & & \\
 & & 0 & & 6 & & 11 & & 6 & & 1 & & \\
 & 0 & & 24 & & 50 & & 35 & & 10 & & 1 & \\
0 & & 120 & & 274 & & 225 & & 85 & & 15 & & 1
\end{array}
$$

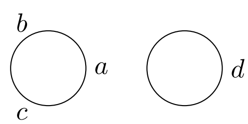
The recurrence (iii) can be proved by observing that

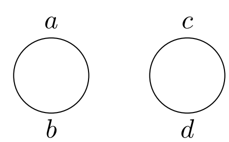
$$(x)_n = (x)_{n-1}(x - n + 1) = x(x)_{n-1} - (n-1)(x)_{n-1},$$

so the coefficient of $x^k$ in $(x)_n$ is the coefficient of $x^{k-1}$ in $(x)_{n-1}$ less $n-1$ times the coefficient of $x^k$ in $(x)_{n-1}$.

Now here is an interesting fact: $|s(n,k)|$ is the number of ways to seat $n$ people around $k$ circular tables with at least one person at each table, where we consider two seatings to be equivalent if everyone has the same left and right neighbor. The particular table one is assigned to is irrelevant; all that matters is the partitioning of the people into tables and the circular ordering around the tables. If you know something about finite permutation groups, it is the number of permutations of $n$ objects with $k$ cycles.

For example, $|s(4,2)| = 11$ is the number of ways to seat four people around two circular tables with no table left empty.




In the left-hand diagram, we have seated three people at one table and one at another. There are $8 = 4 \cdot 2$ ways to do this, four ways to pick the lonely person and two circular permutations of the remaining three people. In the right-hand diagram, we have seated two people at each table. There are three ways to do this, since $a$ can be paired with either $b$, $c$, or $d$. (Switching tables does not matter.) There are $8 + 3 = 11$ seatings in all.

To prove the relationship between circular table seatings and falling factorials, we argue that the numbers of seatings for various values of $n$ and $k$ satisfy the recurrence given by (i′), (ii′), and (iii′). As the recurrence uniquely determines the function, it follows by induction that this number must equal $|s(n,k)|$. Certainly there is only one way to seat $n$ people at $n$ tables with no empty tables, as this can only happen if there is one person at each table, therefore (i′) holds. There are no ways to seat $n$ people at 0 tables, therefore (ii′) holds. Finally, for (iii′), to add an $n$th person to obtain a seating at $k$ tables, we either add a new table with that new person alone to an existing seating of $n-1$ people at $k-1$ tables, giving the first term of (iii′), or we insert that person into an existing seating of $n-1$ people at $k$ tables in any one of $n-1$ possible positions, giving the second term of (iii′).

Here are a few more interesting properties that are not hard to prove.

(iv) $|s(n,1)| = (n-1)!$ for all $n \geq 1$

(v) $\sum_{k=0}^{n} s(n,k) = 0$

(vi) $\sum_{k=0}^{n} |s(n,k)| = n!$

Intuitively, (vi) says that the number of permutations of an $n$-element set is the sum over all $k$ of the number of permutations with $k$ cycles. The property (iv) says that there are $(n-1)!$ permutations of an $n$-element set with a single cycle. An alternative interpretation of (iv) is the number of ways to seat $n$ people at one table—there are $n!$ permutations, but we must divide by $n$ because $n$ permutations give the same circular seating.

## Stirling Numbers of the Second Kind

*Stirling numbers of the second kind* are denoted $S(n,k)$. The number $S(n,k)$ is the number of ways to partition a set of size $n$ into $k$ nonempty sets. Equivalently, $S(n,k)$ is the number of equivalence relations on a set of size $n$.

These numbers also satisfy a Pascal-like recurrence:

(i) $S(n,n) = 1$ for all $n \geq 0$

(ii) $S(n,0) = 0$ for all $n \geq 1$

(iii) $S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k)$ for all $0 < k < n$.

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 0 & & 1 & & & & & \\
 & & & & 0 & & 1 & & 1 & & & & \\
 & & & 0 & & 1 & & 3 & & 1 & & & \\
 & & 0 & & 1 & & 7 & & 6 & & 1 & & \\
 & 0 & & 1 & & 15 & & 25 & & 10 & & 1 & \\
0 & & 1 & & 31 & & 90 & & 65 & & 15 & & 1
\end{array}
$$

Intuitively, there is one equivalence relation on an $n$ element set with $n$ equivalence classes, namely the identity relation; there are no equivalence relations on an $n$-element set with no equivalence classes for $n \geq 1$; and to add a new element and have $k$ classes, one can either add it to an equivalence class of an existing equivalence relation with $k$ classes in $k$ possible ways or add it as a new singleton class to an existing equivalence relation with $k-1$ classes.

There is a summation formula for $S(n,k)$:

$$S(n,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n$$

To prove this, note first that it suffices to prove

$$k! \cdot S(n,k) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n.$$

The left-hand side is the number of surjective functions $f : X \to Y$, where $|X| = n$ and $|Y| = k$. This is because such a function $f$ is determined by the equivalence relation $x \equiv y$ iff $f(x) = f(y)$ on $X$ and an assignment of a value in $Y$ to each equivalence class. There are $S(n,k)$ ways to choose an equivalence relation on $X$ with $k$ equivalence classes and $k!$ ways to assign values in $Y$ to the equivalence classes.

It therefore suffices to prove that the number of surjective functions $X \to Y$, where $|X| = n$ and $|Y| = k$, is

$$\sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n$$

We can do this using the inclusion-exclusion principle. We did not say much in class about the inclusion-exclusion principle, but you have seen small instances of it in the homework for two and three sets:

$$
\begin{aligned}
|A \cup B| &= |A| + |B| - |A \cap B| \\
|A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
\end{aligned}
$$

In general, let $Y = \{1, 2, \ldots, k\}$. Then

$$\begin{aligned} |\bigcup_{i=1}^{k} A_i| &= \sum_{i=1}^{k} |A_i| - \sum_{1 \le i < j \le k} |A_i \cap A_j| + \sum_{1 \le i < j < m \le k} |A_i \cap A_j \cap A_m| - \cdots \pm |A_1 \cap \cdots \cap A_n| \\ &= \sum_{j=1}^{k} (-1)^{j+1} \sum_{\substack{B \subseteq Y \\ |B|=j}} |\bigcap_{i \in B} A_i| \end{aligned}$$

This can be proved by induction on $k$.

To apply this to the problem at hand, let $X = \{1, 2, \ldots, n\}$ and $Y = \{1, 2, \ldots, k\}$. For $i \in Y$, let

$$A_i = \{f : X \to Y \mid \forall x \in X \ f(x) \neq i\}.$$

Then $\bigcup_{i=1}^{k} A_i$ is the set of functions $f : X \to Y$ that are not surjective. Also, for $B \subseteq Y$, $|B| = j$,

$$\bigcap_{i \in B} A_i = \{f : X \to Y \mid \forall x \in X \ f(x) \in Y - B\} \qquad |\bigcap_{i \in B} A_i| = |Y - B|^n = (k - j)^n.$$

Thus

$$\begin{aligned} |\bigcup_{i=1}^{k} A_i| &= \sum_{j=1}^{k} (-1)^{j+1} \sum_{\substack{B \subseteq Y \\ |B|=j}} |\bigcap_{i \in B} A_i| \\ &= \sum_{j=1}^{k} (-1)^{j+1} \sum_{\substack{B \subseteq Y \\ |B|=j}} (k - j)^n \\ &= \sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} (k - j)^n. \end{aligned}$$

The number of surjective functions is the total number of functions minus the number of non-surjective functions, or

$$\begin{aligned} k^n - |\bigcup_{i=1}^{k} A_i| &= k^n - \sum_{j=1}^{k} (-1)^{j+1} \binom{k}{j} (k - j)^n \\ &= k^n + \sum_{j=1}^{k} (-1)^{j} \binom{k}{j} (k - j)^n \\ &= \sum_{j=0}^{k} (-1)^{j} \binom{k}{j} (k - j)^n. \end{aligned}$$

Another interesting property of the Stirling numbers of the second kind is

$$m^n = \sum_{k=0}^{n} S(n, k) P(m, k) = \sum_{k=0}^{n} S(n, k) (m)_k. \tag{1}$$

Intuitively, there are $m^n$ functions from an $n$-element set to an $m$-element set. Each such function $f$ determines an equivalence relation $x \equiv y$ iff $f(x) = f(y)$. We can first choose the equivalence relation on the $n$-element set, then choose the values for the elements of each equivalence class. There are $S(n, k)$ ways to choose an equivalence relation with $k$ equivalence classes in the first step, and there are $P(m, k)$ ways to choose the values for the $k$ equivalence classes in the second step. Thus there are $S(n, k)P(m, k)$ ways to choose a function with $k$ equivalence classes, therefore $\sum_{k=0}^{n} S(n, k)P(m, k)$ functions in all.

## Relationship between Stirling Numbers of the First and Second Kinds

If you have taken linear algebra, you will appreciate this part. The polynomials in one variable form an infinite-dimensional vector space. The usual basis for that space is the set of monomials $1, x, x^2, x^3, \ldots$. The polynomials of degree $m$ or less form a subspace of dimension $m+1$ with basis $1, x, x^2, \ldots, x^m$.

There is another basis for the space of all polynomials, namely $(x)_0, (x)_1, (x)_2, (x)_3, \ldots$. It is also fairly clear that $(x)_0, (x)_1, (x)_2, \ldots, (x)_m$ form a basis for the polynomials of degree $m$. We have already seen that

$$(x)_n = \sum_{k=0}^{n} s(n,k)x^k,$$

so the Stirling numbers of the first kind $s(n,k)$ for $0 \le k \le m$, arranged in an $(m+1) \times (m+1)$ triangular matrix, form a linear transformation that transforms the basis $1, x, x^2, \ldots, x^m$ to the basis $(x)_0, (x)_1, (x)_2, \ldots, (x)_m$.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & 0 & 0 & 0 \\ 0 & -6 & 11 & -6 & 1 & 0 & 0 \\ 0 & 24 & -50 & 35 & -10 & 1 & 0 \\ 0 & -120 & 274 & -225 & 85 & -15 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \\ x^5 \\ x^6 \end{bmatrix} = \begin{bmatrix} (x)_0 \\ (x)_1 \\ (x)_2 \\ (x)_3 \\ (x)_4 \\ (x)_5 \\ (x)_6 \end{bmatrix}$$

Now the Stirling numbers of the second kind transform the space in the inverse direction. The property (1) holds for all $m$, and two polynomials of degree $n$ that agree on $n+1$ inputs agree everywhere, therefore

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

This says that the Stirling numbers of the second kind $S(n,k)$ for $0 \le k \le m$, arranged in an $(m+1)\times(m+1)$ triangular matrix, form a linear transformation that transforms the basis $(x)_0, (x)_1, (x)_2, \ldots, (x)_m$ to the basis $1, x, x^2, \ldots, x^m$.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 1 & 7 & 6 & 1 & 0 & 0 \\ 0 & 1 & 15 & 25 & 10 & 1 & 0 \\ 0 & 1 & 31 & 90 & 65 & 15 & 1 \end{bmatrix} \cdot \begin{bmatrix} (x)_0 \\ (x)_1 \\ (x)_2 \\ (x)_3 \\ (x)_4 \\ (x)_5 \\ (x)_6 \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \\ x^5 \\ x^6 \end{bmatrix}$$

Thus the two matrices are inverses:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & 0 & 0 & 0 \\ 0 & -6 & 11 & -6 & 1 & 0 & 0 \\ 0 & 24 & -50 & 35 & -10 & 1 & 0 \\ 0 & -120 & 274 & -225 & 85 & -15 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 1 & 7 & 6 & 1 & 0 & 0 \\ 0 & 1 & 15 & 25 & 10 & 1 & 0 \\ 0 & 1 & 31 & 90 & 65 & 15 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$