175

L4 **Counting with repetition**

Q How many words of length 4 are in alphabetical order?

Some examples: AABZ, AAAA.

A1 Trick: partition according to # repetitions

All letters different: $\binom{26}{4}$ ← 4-element subsets of $\{A, B, \dots, Z\}$

Two letters same: $\binom{26}{3} \binom{3}{1}$

pick 3 distinct letters. pick one of 3 to duplicate

(Two same) + (other two same): eg AABB: $\binom{26}{2}$

pick 2 distinct letters.

Three same: $2\binom{26}{2}\binom{2}{1}$

pick 2 distinct letters. choose which to triplicate

All same: $\binom{26}{1}$.

choose one letter

Thus answer is $\binom{26}{4} + \binom{26}{3}\binom{3}{1} + \binom{26}{2} + \binom{26}{2}\binom{2}{1} + \binom{26}{1}$

Good: we now have an answer. But it is difficult to generalize to words of arbitrary length.

A2 Trick: vary alphabet size

- Alphabet = {A}.
  #words = 1 (AAAA)

- Alphabet = {A, B}
  Words of length 4:

#A	#B	
0	4	BBBB
1	3	ABBB
2	2	AABB
3	1	AAAB
4	0	AAAA.

$\Rightarrow$ # words = 5

- Alphabet = {A, B, C}.

#A

0		1		2		3		4	
#B	#C	#B	#C	#B	#C	#B	#C	#B	#C
0	4	0	3	0	2	0	1	0	0
1	3	1	2	1	1	1	0		
2	2	2	1	2	0				
3	1	3	0						
4	0								

Thus:

$\#\text{words} = 5+4+3+2+1$

At this point, you might notice a pattern. Letting L = word size, we have:

alphabet	#words of length L
$\{A\}$	1
$\{A,B\}$	$\frac{L+1}{1}$
$\{A,B,C\}$	$\sum_{k=1}^{L+1} k = \frac{(L+1)(L+2)}{2\cdot 1}$

This might lead you to predict that the next row of the table is:

alphabet	#words of length L
$\{A,B,C,D\}$	$\frac{(L+1)(L+2)\cdot(L+3)}{3\cdot 2\cdot 1}$

Let's check:

#A	:	0	1	2	3	4
#slots for B, C, D	:	4	3	2	1	0
#words from from $\{B,C,D\}$ (see above)	:	$\frac{(4+1)(4+2)}{2}$	$\frac{(3+1)(3+2)}{2}$	$\frac{(2+1)(2+2)}{2}$	$\frac{(1+1)(1+2)}{2}$	$\frac{(0+1)(0+2)}{2}$

Thus total # words is:

$$\sum_{i=0}^{L} \frac{(i+1)(i+2)}{2} \qquad (*)$$

Using the facts that

$$\sum_{i=0}^{L} i^2 = \frac{L(L+1)(2L+1)}{6}, \quad \sum_{i=0}^{L} i = \frac{L(L+1)}{2}$$

we may show that:

$$(*) = \frac{(L+1)(L+2)(L+3)}{3.2.1.}$$

Thus we have proved the conjecture.

You may have noticed that we computed the answer for an alphabet of size $M+1$ by using the answer for an alphabet of size M. We can formalize this fact as follows:

Suppose # words of length L from an alphabet of size M is

$$\binom{L+(M-1)}{M-1}.$$

[This formula reduce to the answers we found for $M=1, 2, 3, 4$].

Then, can we show that # words using alphabet of size $M+1$ is:

$$\binom{L+(M+1)-1}{(M+1)-1} = \binom{L+M}{M} \;? \qquad (**)$$

The answer, of course, is yes, and here is how do it. Let alphabet be $\{A_1, A_2, \ldots A_{M+1}\}$. Then, we have the table:

$\#A_1$ in word:	0	1	2	$\ldots$	L
# slots for letters in $\{A_2, \ldots, A_{M+1}\}$	L	$L-1$	$L-2$	$\ldots$	0
# words from $\underbrace{\{A_2, \ldots, A_{M+1}\}}_{\text{size}=M}$:	$\binom{L+(M-1)}{M-1}$	$\binom{(L-1)+(M-1)}{M-1}$	$\binom{(L-2)+(M-1)}{M-1}$		$\binom{0+(M-1)}{M-1}$

Answer is given by sum of elements in bottom row of table:

$$\sum_{j=0}^{L} \binom{j+(M-1)}{M-1} \qquad (1)$$

Let: $\quad m = j + (M-1)$

$\quad\quad\quad k = M-1$

Then

$$(1) = \sum_{m=k}^{L+k} \binom{m}{k} \qquad (2)$$

Further, let:

$$n = L + k$$

Then

$$(2) = \sum_{m=k}^{n} \binom{m}{k}$$

There is a well-known identity:

$$\sum_{m=k}^{n} \binom{m}{k} = \binom{n+1}{k+1} \qquad (3)$$

or, in terms of our parametrization,

$$(3) = \binom{L+k+1}{k+1} = \binom{L+M}{M}$$

Thus we have established ($\ast\ast$).

Optional: We now have all the results in place to prove that # words of length L from an alphabet of size M is

$$\binom{L+M-1}{M-1}$$

for any L, M. The way we do it is via induction. To see how, tune in later in the course!

A3. This problem is actually quite easy to solve once we phrase it in the following way:

Given L objects (representing the letters), how many ways are there of dividing them into M groups?

Example:

$$\text{alphabet} = \{A, B, C\} \quad (M = 3)$$

$$\text{word} = ABBC \quad (L = 4)$$

Think of a word as a string of "stars" seperated by "bars".

$$ABBC = \underset{A}{*} \;\Big|\; \underset{B}{*\,*} \;\Big|\; \underset{C}{*}$$

- We assigned 1 star to first group, 2 to second group, 1 to third group.

Notice: # bars = 1 less than alphabet size.

More examples:

BBCC = $\underset{A}{} \mid \underset{B}{* *} \mid \underset{C}{* *}$

CCCC = $\underset{A}{} \mid \underset{B}{} \mid \underset{C}{* * * *}$

Clearly the positions of the bars relative to the stars is important here. This suggests that we should treat the bars on the same footing as the stars:

BBCC = $\underline{|}\ \underline{*}\ \underline{*}\ \underline{|}\ \underline{*}\ \underline{*}$ # slots = 4 + 2

CCCC = $\underline{|}\ \underline{|}\ \underline{*}\ \underline{*}\ \underline{*}\ \underline{*}$

Now we see that our problem is really that of counting the # ways of placing $(M-1)$ bars into $(L+(M-1))$ slots:

As before (cf L3, p5) the answer is

$$\binom{L+(M-1)}{M-1}$$

Concretely, for $L=4, M=3$, there are 6 slots into which I can place the 1st bar. Having placed the 1st bar, there are 5 slots remaining to position the 2nd bar.

In all there seems to be:

$$6 \cdot 5$$

ways to place 2 bars in 6 slots. Thats not quite right because the bars are indistinguishable, so we need to divide by the # orderings of 2 objects, which is $2 \cdot 1$.

Thus:

$$\# \text{ words} = \frac{6 \cdot 5}{2 \cdot 1} = \frac{6!}{(6-2)! \cdot 2!} = \binom{6}{2}$$

How does the result,

$$\begin{matrix}\# \text{words of length } L \\ \text{in alphabet of size } M\end{matrix} = \binom{L+(M-1)}{M-1}$$

compare to approach 1?

$L = 4,\ M = 26 \Rightarrow$

$$\#\text{words} = \underbrace{\binom{4+(26-1)}{26-1}}_{\text{choose bars.}} = \underbrace{\binom{(26-1)+4}{4}}_{\text{choose stars.}} = \binom{29}{4}$$

$$= 23{,}751$$

$$= \binom{26}{4} + \binom{26}{3}\binom{3}{1} + \binom{26}{2} + \binom{26}{2}\binom{2}{1} + \binom{26}{1}$$

Anther way to phrase our result is:

Lemma: The # k-multisets of an n-set is

$$\binom{k+(n-1)}{n-1} = \binom{n-1+k}{k}$$