Problem 1. How many n-letter words coming from $[m]$ use the letter 1 at least twice and each other letter at least once?

Let B_1 be the collection of words in $[m]^n$ that use 1 at most once and for $i \in [2, m]$, let B_i be the collection of words in $[m]^n$ that do not use i. It is easy to observe that $\bigcap_{i\in[m]} B_i^C$ is precisely the set that we wish to count. For an index set $S \subseteq [m]$, we find that

$$\left|\bigcap_{i\in S} B_i\right| = \begin{cases} (m-|S|)^n & \text{if } 1 \notin S; \\ (m-|S|)^n + n(m-|S|)^{n-1} & \text{if } 1 \in S. \end{cases}$$

Therefore, by inclusion-exclusion, the number of words satisfying the desired property is

$$\begin{aligned}\sum_{S\subseteq[m]} (-1)^{|S|} \left|\bigcap_{i\in S} B_i\right| &= \sum_{\substack{S\subseteq[m]:\\ 1\notin S}} (-1)^{|S|}(m-|S|)^n + \sum_{\substack{S\subseteq[m]:\\ 1\in S}} (-1)^{|S|}\left((m-|S|)^n + n(m-|S|)^{n-1}\right) \\ &= \sum_{S\subseteq[m]} (-1)^{|S|}(m-|S|)^n + \sum_{\substack{S\subseteq[m]:\\ 1\in S}} (-1)^{|S|} n(m-|S|)^{n-1} \\ &= \sum_{i=0}^{m} \binom{m}{i}(-1)^i(m-i)^n + n\sum_{i=1}^{m}\binom{m-1}{i-1}(-1)^i(m-i)^{n-1}.\end{aligned}$$

We now discuss an alternative proof for the number of derangements.

Theorem 2. *Let D_n be the set of derangements of $[n]$, then $|D_n| = \sum_{k=0}^{n}(-1)^k\frac{n!}{k!}$.*

Proof. For convenience, let $[n]_k$ denote the set of k-letter words coming from $[n]$ such that no letter is repeated (hence $|[n]_k| = n!/(n-k)!$). Define the function σ by $\sigma(w) = (-1)^{n-k}$ if $w \in [n]_k$. Also, for a word w, let $F(w)$ be the smallest i such that *either* i does not appear in w *or* i is fixed in w (that is, the ith letter in w is i). Of course, F is not defined for all words in $\bigcup_{k=0}^{n}[n]_k$; but it is easy to observe that $\{w \in \bigcup_{k=0}^{n}[n]_k : F(w) \text{ undefined}\} = D_n$.

Now, consider the collection of words for which F is defined. Suppose that w is such that $F(w)$ does not appear in w, then define $s(w)$ to be the word formed by inserting $F(w)$ into its fixed position. Similarly, if w is such that $F(w)$ appears in w, then define $s(w)$ to be the word formed by removing $F(w)$ from w. For example $s(1354) = 354$ and $s(21456) = 213456$. Notice that w and $s(w)$ always differ in length by *exactly* one (so $\sigma(w) = -\sigma(s(w))$) and that $s(s(w)) = w$ (i.e. it is an involution).

Define $Q_n = \sum_{w\in\bigcup_{k=0}^{n}[n]_k}\sigma(w)$; we will rewrite Q_n in two different ways. Firstly,

$$Q_n = \sum_{k=0}^{n}(-1)^{n-k}|[n]_k| = \sum_{k=0}^{n}(-1)^{n-k}\frac{n!}{(n-k)!} = \sum_{k=0}^{n}(-1)^k\frac{n!}{k!}.$$

On the other hand, we can pair up words based on the involution s, so

$$Q_n = \sum_{\substack{w:\\ F(w)\text{ undefined}}} \sigma(w) + \sum_{\{w,s(w)\}} (\sigma(w) + \sigma(s(w)) = |D_n|.$$

Hence, $|D_n| = Q_n = \sum_{k=0}^{n}(-1)^k\frac{n!}{k!}$. □

Here is another proof of the principle of inclusion-exclusion using the same technique as above. In other words, this technique can be used to approach any inclusion-exclusion-type problem.

Theorem 3. *Let U be a finite set and let $B_1, \dots, B_n \subseteq U$. Then*

$$\left|U \setminus \bigcup_{i=1}^{n} B_i\right| = \sum_{S\subseteq[n]} (-1)^{|S|} \left|\bigcap_{i\in S} B_i\right|.$$

Proof. Consider the following set $\mathcal{B} = \{(x,S) \in U \times [n] : x \in \bigcap_{i\in S} B_i\}$. In other words, $\mathcal{B}$ consists of pairs of an index set S and an element of U that satisfies at least all the events indexed by S. Obviously, if we fix a set $T \subseteq [n]$, then $|\{x \in U : (x,T) \in \mathcal{B}\}| = \left|\bigcap_{i\in T} B_i\right|$. For a pair $(x,S) \in \mathcal{B}$, define $\sigma(x,S) = (-1)^{|S|}$. Also, for any $x \in U$, define $F(x) = \min\{i \in [n] : x \in B_i\}$ and note that $F(x)$ is undefined if and only if $x \notin B_i$ for all $i \in [n]$ (i.e. $x \in U \setminus \bigcup_{i=1}^{n} B_i$). Additionally note that if $x \in U \setminus \bigcup_{i=1}^{n} B_i$, then $(x,S) \in \mathcal{B}$ if and only if $S = \emptyset$.

Now, if $F(x)$ is defined, for any set index set $S \subseteq [n]$, let $s(x,S) = (x, S \triangle F(x))$. Notice that as $x \in B_{F(x)}$ (by definition), we have that $(x, S\triangle F(x)) \in \mathcal{B}$. Further, s is an involution (as the symmetric difference is an involution) and $\sigma(x,S) = -\sigma(s(x,S))$.

Let $Q = \sum_{(x,S)\in\mathcal{B}} \sigma(x,S)$; we will rewrite Q in two ways. Firstly,

$$Q = \sum_{S\subseteq[n]} (-1)^{|S|} |\{x : (x,S) \in \mathcal{B}\}| = \sum_{S\subseteq[n]} (-1)^{|S|} \left|\bigcap_{i\in S} B_i\right|.$$

On the other hand, we can pair up elements of $\mathcal{B}$ based on the involution s, so

$$Q = \sum_{\substack{(x,S)\in\mathcal{B}:\\ F(x)\text{ undefined}}} \sigma(x,S) + \sum_{\{(x,S),s(x,S)\}} (\sigma(x,S) + \sigma(s(x,S))) = \left|U \setminus \bigcup_{i=1}^{n} B_i\right|.$$

□

Personally, I prefer to use this matching-type argument over inclusion-exclusion as, in my opinion, it is much more elegant. If you would like to practice this technique, here are a couple things to try to prove.

(1) For $m \le n$, $\sum_{i=0}^{m} (-1)^i \binom{n}{i} = (-1)^m \binom{n-1}{m}$.

(2) For $k < n$, $\sum_{i=0}^{n} (-1)^i \binom{n}{i}\binom{i}{k} = 0$.

(3) The number of surjections from $[n] \to [m]$ is precisely $\sum_{i=0}^{m} (-1)^i \binom{m}{i} (m-i)^n$. Hint: consider pairs (X,Y) where $X \subseteq [m]$ and $Y \in ([m] \setminus X)^n$.