Inclusion - Exclusion Formula

We have seen that

$$\begin{aligned} P(A_1 \cup A_2) = {} & P(A_1) + P(A_2) \quad \text{inclusion} \\ & - P(A_1 \cap A_2) \quad \text{exclusion} \end{aligned}$$

and

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) \quad \text{inclusion} \\ & - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \quad \text{exclusion} \\ & + P(A_1 \cap A_2 \cap A_3) \quad \text{inclusion} \end{aligned}$$

We can see the pattern. In general, we have the following result:

Inclusion-Exclusion formula

Let J_n be a sorted subset of the set $\{1, 2, 3, ..., n\}$. We write $|J_n|$ to denote the number of elements in J_n. For example, if $n = 3$

$$\begin{aligned} |J_3| = 1 &\Rightarrow J_3 = \{1\}, \{2\}, \text{ or } \{3\} \\ |J_3| = 2 &\Rightarrow J_3 = \{1, 2\}, \{1, 3\}, \text{ or } \{2, 3\} \\ |J_3| = 3 &\Rightarrow J_3 = \{1, 2, 3\} \end{aligned}$$

and if $n = 4$

$$\begin{aligned} |J_4| = 1 &\Rightarrow J_4 = \{1\}, \{2\}, \{3\} \text{ or } \{4\} \\ |J_4| = 2 &\Rightarrow J_4 = \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\} \text{ or } \{3, 4\} \\ |J_4| = 3 &\Rightarrow J_4 = \{1, 2, 3\}, \{1, 2, 4\} \text{ or } \{2, 3, 4\} \\ |J_4| = 3 &\Rightarrow J_4 = \{1, 2, 3, 4\} \end{aligned}$$

Using this notation, the inclusion-exclusion formula can be written as:

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{j=1}^{n} (-1)^{j-1} \sum_{|J_n|=j} P\left(\cap_{i\in J_n} A_i\right)$$

$$\begin{aligned}
&= \sum_{1\le i\le n} P\left(A_i\right) \quad \text{inclusion} \\
&- \sum_{1\le i<j\le n} P\left(A_i \cap A_j\right) \quad \text{exclusion} \\
&+ \sum_{1\le i<j<k\le n} P\left(A_i \cap A_j \cap A_k\right) \quad \text{inclusion} \\
&- \sum_{1\le i<j<k<l\le n} P\left(A_i \cap A_j \cap A_k \cap A_l\right) \quad \text{exclusion} \\
&\vdots \\
&+ (1)^{n-1} P\left(A_1 \cap A_2 \cdots \cap A_n\right)
\end{aligned}$$

For example, if $n = 4$

$$P\left(\cup_{i=1}^{4} A_i\right) = \sum_{j=1}^{4} (-1)^{j-1} \sum_{|J_4|=j} P\left(\cap_{i\in J_4} A_i\right)$$

$$= P\left(A_1\right) + P\left(A_2\right) + P\left(A_3\right) + P\left(A_4\right) \quad \text{inclusion}$$

$$- \left[P\left(A_1 \cap A_2\right) + P\left(A_1 \cap A_3\right) + P\left(A_1 \cap A_4\right) + P\left(A_2 \cap A_3\right) + P\left(A_2 \cap A_4\right) + P\left(A_3 \cap A_4\right)\right] \quad \text{exclusion}$$

$$+ \left[P\left(A_2 \cap A_3 \cap A_4\right) + P\left(A_1 \cap A_3 \cap A_4\right) + P\left(A_1 \cap A_2 \cap A_4\right) + P\left(A_1 \cap A_2 \cap A_3\right)\right] \quad \text{inclusion}$$

$$- P\left(A_1 \cap A_2 \cap A_3 \cap A_4\right) \quad \text{exclusion}$$

Proof:

By induction. The result clearly holds for $n = 1$

Suppose that the result holds for $n = k > 1$. We will show that in such case the result also holds for $n = k + 1$. In fact,

$$P\left(\cup_{i=1}^{k+1} A_i\right) = P\left(\left[\cup_{i=1}^{k} A_i\right] \cup A_{k+1}\right) = P\left(\cup_{i=1}^{k} A_i\right) + P\left(A_{k+1}\right) - P\left(\left[\cup_{i=1}^{k} A_i\right] \cap A_{k+1}\right)$$

$$= \sum_{j=1}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\cap_{i\in J_k} A_i\right) + P\left(A_{k+1}\right) - P\left(\cup_{i=1}^{k}\left[A_i \cap A_{k+1}\right]\right)$$

$$= \sum_{j=1}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\cap_{i\in J_k} A_i\right) + P\left(A_{k+1}\right) - \sum_{j=1}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\left[\cap_{i\in J} A_i\right] \cap A_{k+1}\right)$$

$$= \sum_{i=1}^{k+1} P\left(A_i\right) + \sum_{j=2}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\cap_{i\in J_k} A_i\right)$$
$$- \sum_{j=1}^{k-1} (-1)^{j-1} \sum_{|J_k|=j} P\left(\left[\cap_{i\in J_k} A_i\right] \cap A_{k+1}\right) + (-1)(-1)^{k-1} P\left(A_1 \cap \cdots \cap A_{k+1}\right)$$

$$= \sum_{j=1}^{k+1} (-1)^{j-1} \sum_{|J_{k+1}|=j} P\left(\cap_{i\in J} A_i\right)$$

The last equality holds because

$$\sum_{j=2}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\cap_{i\in J_k} A_i\right) - \sum_{j=1}^{k-1} (-1)^{j-1} \sum_{|J_k|=j} P\left(\left[\cap_{i\in J_k} A_i\right] \cap A_{k+1}\right) =$$

$$= \sum_{j=2}^{k} (-1)^{j-1} \sum_{|J_k|=j} P\left(\cap_{i\in J_k} A_i\right) + \sum_{j=1}^{k-1} (-1)^{j} \sum_{|J_k|=j} P\left(\left[\cap_{i\in J_k} A_i\right] \cap A_{k+1}\right) =$$

$$= \sum_{j=2}^{k} (-1)^{j-1} \sum_{|J_{k+1}|=j} P\left(\cap_{i\in J_k} A_i\right)$$