

MATH 203: PRINCIPLES OF STATISTICS 1
FALL 2024

PROOF OF THE GENERAL ADDITION FORMULA

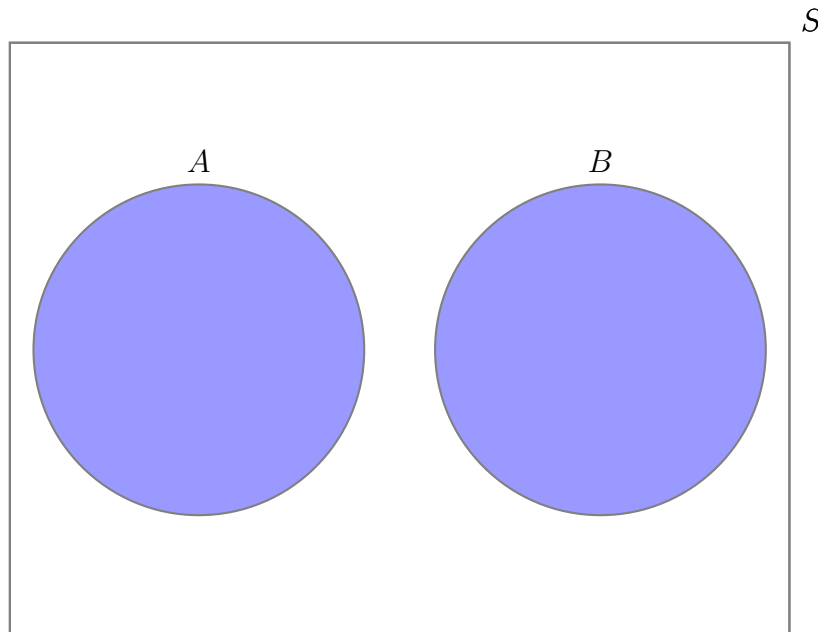
Probability Rule III says that for two events A and B in sample space S such that

$$A \cap B = \emptyset$$

we have

$$P(A \cup B) = P(A) + P(B).$$

Essentially, this says that the probability ‘content’ of the set $A \cup B$ can be computed by simply adding the two probability contents of A and B . The associated Venn diagram would look like this: we simply add up the (blue) probability content



We can see immediately that this result extends to any number of mutually exclusive events: that is, if $A_1, A_2, \dots, A_K$ are such that for all j and k with $j \neq k$

$$A_j \cap A_k = \emptyset \quad j, k = 1, 2, \dots, K$$

(that is, none of the A_k s overlap each other), then

$$P(A_1 \cup A_2 \cup \dots \cup A_K) = P(A_1) + P(A_2) + \dots + P(A_K)$$

or more concisely

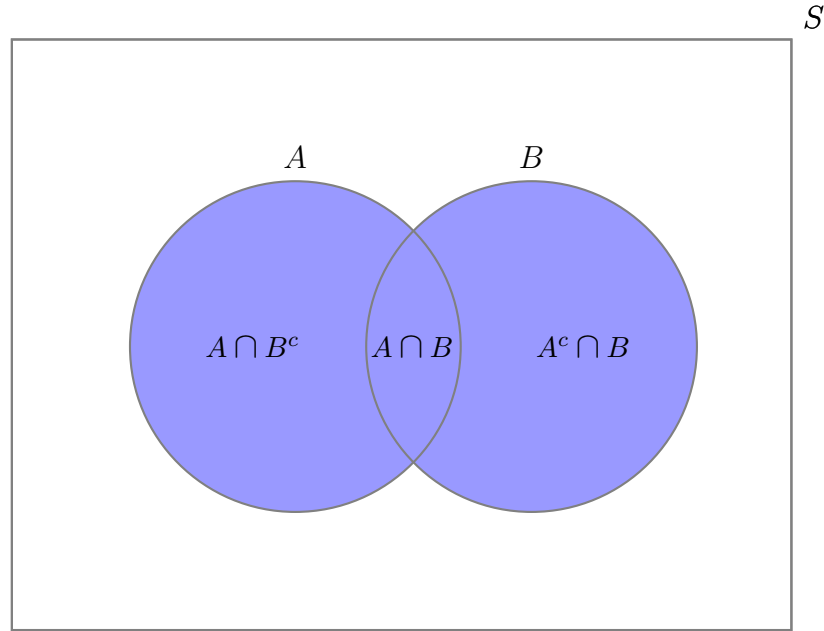
$$P\left(\bigcup_{k=1}^K A_k\right) = \sum_{k=1}^K P(A_k).$$

In fact, this result holds even if we allow K to become infinite, although to see that we need to assume a little more structure concerning S .

The *General Addition Rule* extends this result: for any two events A and B , we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

To see why, consider the following Venn diagram



The first key realization is that we now can use a partitioning argument, and write (as indicated by the Venn diagram) that

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)$$

that is, the blue shaded region overall can be broken down into the three non-overlapping portions. Hence by Rule III, we have

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B).$$

Our next task is to compute $P(A \cap B^c)$, which we can do by using another partitioning argument now for A , that is,

$$A = (A \cap B) \cup (A \cap B^c)$$

that is, A can be broken down into two non-overlapping parts,

- $A \cap B$: elements that are in A and in B ;
- $A \cap B^c$: elements that are in A and **not** in B .

Hence, again by Rule III, we have

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

so that on rearrangement, we obtain

$$P(A \cap B^c) = P(A) - P(A \cap B).$$

By an identical argument for B , we have that

$$P(A^c \cap B) = P(B) - P(A \cap B).$$

Now we are in a position to obtain the final result:

$$\begin{aligned}
 P(A \cup B) &= P(A \cap B^c) + P(A \cap B) + P(A \cap B^c) \\
 &= (P(A) - P(A \cap B)) + P(A \cap B) + (P(B) - P(A \cap B)) \\
 &= P(A) + P(B) - P(A \cap B).
 \end{aligned}$$

For more than two events, we can simply replicate the above arguments. For three events

$$P(A_1 \cup A_2 \cup A_3)$$

first define for simplicity of notation

$$B = A_2 \cup A_3$$

so that by the General Addition Rule

$$P(B) = P(A_2) + P(A_3) - P(A_2 \cap A_3).$$

Then

$$\begin{aligned}
 P(A_1 \cup A_2 \cup A_3) &= P(A_1 \cup B) && \text{properties of union} \\
 &= P(A_1) + P(B) - P(A_1 \cap B) && \text{General Addition Rule} \\
 &= P(A_1) + (P(A_2) + P(A_3) - P(A_2 \cap A_3)) - P(A_1 \cap B) \\
 &= P(A_1) + P(A_2) + P(A_3) - P(A_2 \cap A_3) - P(A_1 \cap B) && (1)
 \end{aligned}$$

For $P(A_1 \cap B)$, we have

$$P(A_1 \cap B) = P(A_1 \cap (A_2 \cup A_3)) = P((A_1 \cap A_2) \cup (A_1 \cap A_3))$$

by the ‘distributivity’ property of union and intersection. Let

$$C_1 = (A_1 \cap A_2) \quad C_2 = (A_1 \cap A_3)$$

Then

$$\begin{aligned}
 P((A_1 \cap A_2) \cup (A_1 \cap A_3)) &= P(C_1 \cup C_2) \\
 &= P(C_1) + P(C_2) - P(C_1 \cap C_2) && \text{Gen. Add. Rule} \\
 &= P(A_1 \cap A_2) + P(A_1 \cap A_3) - P((A_1 \cap A_2) \cap (A_1 \cap A_3)) \\
 &= P(A_1 \cap A_2) + P(A_1 \cap A_3) - P(A_1 \cap A_2 \cap A_3) && \text{prop. of inters...} \\
 &&& (2)
 \end{aligned}$$

Now if we substitute (2) into (1), we obtain

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

as required. The proof can be generalized to any number of events using the principle of mathematical induction.