MATH 323: PROBABILITY
THE CHAIN RULE FOR MULTIPLE EVENTS

For arbitrary events $A_1, A_2, \ldots, A_K$ in sample space S, we wish to compute the probability

$$P(A_1 \cap A_2 \cap \cdots \cap A_K).$$

We can rewrite this using the (trivial !) identity

$$P(A_1 \cap A_2 \cap \cdots \cap A_K) = \frac{P(A_1)}{P(A_1)} \frac{P(A_1 \cap A_2)}{P(A_1 \cap A_2)} \times \cdots \times \frac{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})}{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})} \times P(A_1 \cap A_2 \cap \cdots \cap A_K)$$

where for the leading $K-1$ factors

$$\frac{P(A_1)}{P(A_1)} = \frac{P(A_1 \cap A_2)}{P(A_1 \cap A_2)} = \cdots = \frac{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})}{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})} = 1.$$

Now we imagine 'sliding' the denominators one place to the right to yield

$$P(A_1 \cap A_2 \cap \cdots \cap A_K) = P(A_1) \frac{P(A_1 \cap A_2)}{P(A_1)} \times \cdots \times \frac{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})}{P(A_1 \cap A_2 \cap \cdots \cap A_{K-2})} \times \frac{P(A_1 \cap A_2 \cap \cdots \cap A_K)}{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})}.$$

Then

$$\frac{P(A_1 \cap A_2)}{P(A_1)} = P(A_2|A_1) \tag{1}$$

$$\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} = P(A_3|A_1 \cap A_2) \tag{2}$$

$$\vdots$$

$$\frac{P(A_1 \cap A_2 \cap \cdots \cap A_K)}{P(A_1 \cap A_2 \cap \cdots \cap A_{K-1})} = P(A_K|A_1 \cap A_2 \cap \cdots \cap A_{K-1}) \tag{3}$$

Recall the conditional probability definition: for $P(B) > 0$,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

which holds for arbitrary A and B in S.

- For (1), this is the conditional probability definition with A set to A_2 and B set to A_1;
- for (2), this is the conditional probability definition with A set to A_3 and B set to $A_1 \cap A_2$
- for (3), this is the conditional probability definition with A set to A_K and B set to

$$A_1 \cap A_2 \cap \cdots \cap A_{K-1}.$$

Therefore we have finally that

$$P(A_1 \cap A_2 \cap \cdots \cap A_K) = P(A_1) \times P(A_2|A_1) \cdots \times P(A_K|A_1 \cap A_2 \cap \cdots \cap A_{K-1})$$

which is the *general multiplication* or *chain* rule for probabilities.

CONDITIONAL INDEPENDENCE

We say that A_1 is *conditionally independent* of A_2 ***given*** B if

$$P(A_1|A_2 \cap B) = P(A_1|B) \tag{4}$$

provided both conditional probabilities exist (that is, provided $P(A_2 \cap B) > 0$). Note that by definition

$$P(A_1|A_2 \cap B) = \frac{P(A_1 \cap A_2 \cap B)}{P(A_2 \cap B)} = \frac{P(A_1 \cap A_2|B)P(B)}{P(A_2|B)P(B)} = \frac{P(A_1 \cap A_2|B)}{P(A_2|B)}$$

– note that $P(B) \geq P(A_2 \cap B) > 0$ by assumption. Then from (4)

$$\begin{aligned} P(A_1|A_2 \cap B) = P(A_1|B) \quad &\iff \quad \frac{P(A_1 \cap A_2|B)}{P(A_2|B)} = P(A_1|B) \\ &\iff \quad P(A_1 \cap A_2|B) = P(A_1|B)P(A_2|B) \end{aligned}$$

and we have two equivalent ways of expressing conditional independence.

Note that in this formulation, B is completely arbitrary (apart from the requirement $P(A_2 \cap B) > 0$), and so B could be any 'compound' event, for example $B = A_3$ or

$$B = A_3 \cap A_4$$

or

$$B = (A_3 \cup A_4') \cap A_5$$

etc.