So far, we've looked at 'a priori' probabilities; reasoning about probabilities of subsequent events given data about earlier ones. However, sometimes we want to argue in reverse; reasoning 'a posteriori' about earlier events given data about later ones.

Suppose $\mathcal{U} = \bigcup_{i=1}^{r} A_i$ and B any event with $p(B) > 0$.

Then $B = \bigcup_{i=1}^{r} (B \cap A_i)$ and

$$p(B) = \sum_1^r p(B \cap A_i) = \sum_1^r p(B | A_i)\, p(A_i)$$

Now, for each j, $p(A_j | B) = \dfrac{p(A_j \cap B)}{p(B)}$, so then

$$p(A_j | B) = \frac{p(B | A_j)\, p(A_j)}{\sum_{i=1}^{r} p(B | A_i)\, p(A_i)}$$

This is Bayes' formula, and the study of such things is often called Baysian probability.

Examples

- Have buckets A_1 with 2 red balls and 4 white balls
  A_2 — 1 — — — 2 — —
  A_3 — 5 — — — 4 — — .

  Suppose $p(A_1) = 1/3$, $p(A_2) = 1/6$, and $p(A_3) = 1/2$.
  Event: Pick a box then pick a ball.

  $$p(\text{red}) = p(\text{red} | A_1)\, p(A_1) + p(\text{red} | A_2)\, p(A_2) + p(\text{red} | A_3)\, p(A_3)$$
  $$= 2/6 \cdot 1/3 + 1/3 \cdot 1/6 + 5/9 \cdot 1/2 = 4/9 .$$

  Now suppose the outcome is picking a red ball. Given that, what is $p(A_1)$?

  $$p(A_1 | \text{red}) = \frac{p(A_1 \cap \text{red})}{p(\text{red})} = \frac{p(\text{red} | A_1)\, p(A_1)}{p(\text{red})} = \frac{2/6 \cdot 1/3}{4/9} = \frac{1}{4}.$$