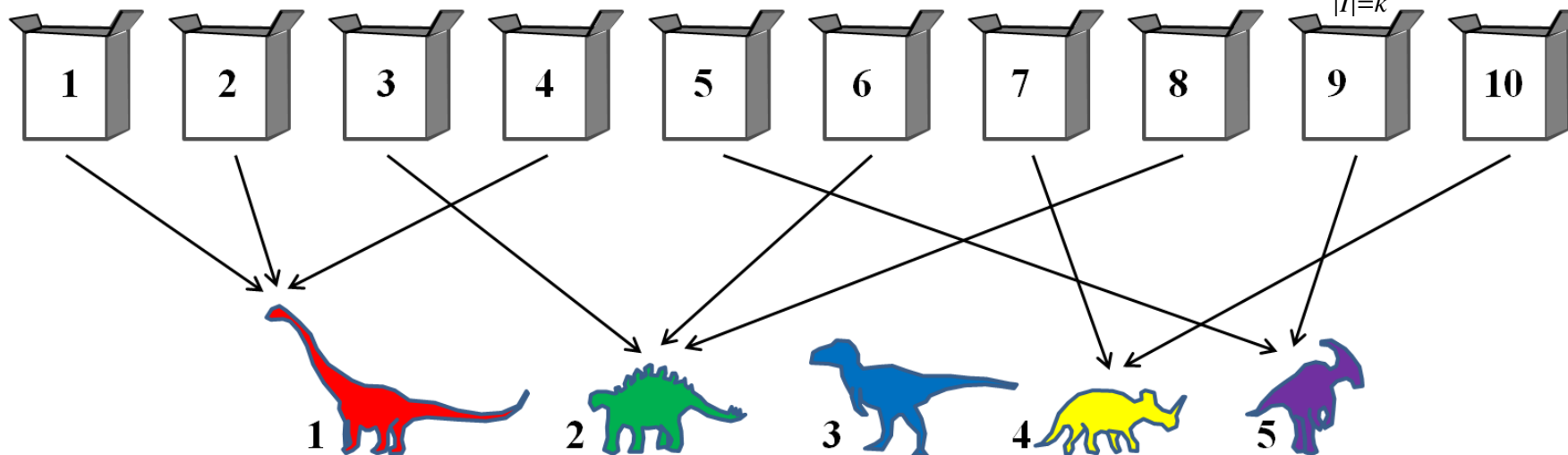




THEOREM OF THE DAY

The Inclusion-Exclusion Principle *If* A_1, A_2, \dots, A_n *are subsets of a set then*

$$|A_1 \cup A_2 \cup \dots \cup A_n| = |A_1| + |A_2| + \dots + |A_n| \\ - (|A_1 \cap A_2| + |A_1 \cap A_3| + \dots + |A_{n-1} \cap A_n|) \\ + (|A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \dots + |A_{n-2} \cap A_{n-1} \cap A_n|) \\ \dots + (-1)^{n-1} |A_1 \cap A_2 \cap \dots \cap A_{n-1} \cap A_n| = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{I \subseteq [n] \\ |I|=k}} |\cap_{i \in I} A_i| \quad ([n] = \{1, 2, \dots, n\}).$$



We illustrate with the classic **Coupon Collector's Problem**: Suppose that we have n different dinosaur models to collect from cereal packets each containing one dinosaur selected uniformly at random from n . What is the probability that we will have all n dinosaurs after munching our way through m packets of cereals?

Putting dinosaurs into packets is modelled as a function from packets to dinosaurs. We have all dinosaurs if the function is **surjective** (aka **onto**: every dinosaur gets at least one packet mapped to it). The illustration above shows a non-surjective sequence of breakfasts: we are still waiting for a packet to map to T. rex.

There are n^m functions from $[m]$ to $[n]$, so our probability is (number of surjections)/ n^m . Let A_i be the subset of functions which *fail* to map to i . Then $|A_1 \cup A_2 \cup \dots \cup A_n|$ counts the number of functions which fail to be surjections. This means that number of surjections = $n^m - |A_1 \cup A_2 \cup \dots \cup A_n|$. Now the number of functions which do not map to $I \subset [n]$ is $|\cap_{i \in I} A_i|$. This is the same as the number of functions $[m] \rightarrow [n - |I|]$, so $|\cap_{i \in I} A_i| = (n - |I|)^m$. We are ready to apply Inclusion-Exclusion:

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{I \subseteq [n] \\ |I|=k}} |\cap_{i \in I} A_i| = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{I \subseteq [n] \\ |I|=k}} (n - |I|)^m = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} (n - k)^m = \sum_{k=1}^{n-1} (-1)^{k-1} \binom{n}{k} (n - k)^m \quad (\text{since summand is zero for } k = n).$$

So **number of surjections** = $n^m - |A_1 \cup A_2 \cup \dots \cup A_n| = n^m - \sum_{k=1}^{n-1} (-1)^{k-1} \binom{n}{k} (n - k)^m = \binom{n}{0} (n - 0)^m + \sum_{k=1}^{n-1} (-1)^k \binom{n}{k} (n - k)^m = \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n - k)^m.$

... and that's just 10 packs for a BETTER THAN 50% chance of getting ALL FIVE, folks!

This principle, almost self-evident but immensely versatile, traces back to Abraham De Moivre in 1718.

Weblink: www.gresham.ac.uk/lectures-and-events/the-mathematics-that-counts.

Further reading: *The Doctrine of Chances: Probabilistic Aspects of Gambling* by Stewart N. Ethier, Springer-Verlag, 2010.

