Coupon Collector Calculator (CCC)

Let $X_1, X_2, \dots$ be i.i.d. random variables taking values in $\{1, \dots, N\}$. Let $p_j = P(X_1 = j)$. Fix $n \in \mathbb{N}$, nonempty $S \subset \{1, \dots, N\}$, and $k \in \{1, \dots, |S|\}$. The main purpose of CCC is to calculate

$$P(|\{X_1, \dots, X_n\} \cap S| \geq k).$$

To use CCC, save the script, `CCC.R`, and `README.md` to the same folder, then run the script in R.

`CCC.R`

When you run this script, it will load two functions, `coup` and `coupx`, which are described below.

`coup`

Let `p` be a vector of nonnegative real numbers with length N. If they do not add up to one, then `coup` will normalize the vector to be a probability vector.

Let $S \subset \{1, \dots, N\}$. Let `g` be the vector of length $|S|$ whose components are the elements of S, listed in ascending order.

Let `n` be any positive integer.

Then `coup(p,g,n)` returns $P(\{X_1, \dots, X_n\} = S)$. This function utilizes the formula described in this paper.

`coupx`

Let `p`, `g`, and `n` be as above. Let `k` be any positive integer not greater than `length(g)`. Then `coupx(p,g,k,n)` returns

$$P(|\{X_1, \dots, X_n\} \cap S| \geq k).$$