Theory of belief functions: Application to machine learning and statistical inference
Exercises on prediction based on multinomial data

Thierry Denœux

Summer 2023

For this exercise, you will need the following R packages: `DescTools`, `evclust`, `ibelief`, `lpSolve`.

1. The calculation of Goodman's simultaneous confidence intervals is implemented in function `MultinomCI` in package `DescTools`. Estimate the coverage rate of these intervals for different values of K (number of levels of the multinomial random variables) and $1 - \alpha$ (confidence levels). To simulate multinomial data, use function `rmultinom`.

2. Write a function that computes the predictive mass and belief functions for $K \leq 3$.

3. By simulation, verify experimentally that requirements R_1 and R_2 in the PBF paper ("Constructing belief functions from sample data using multinomial confidence regions", IJAR, 2006) are satisfied, for $K = 3$.

4. Complete the previous function for the case $K > 3$. (Hint: use function `makeF` of package `evclust` to generate the focal sets, and function `lp` of package `DescTools` for the linear programming). Compare your results to those reported in the PBF paper for the psychiatric patients dataset.

5. Verify by simulation that requirements R_1 and R_2 are satisfied for $K > 3$.

6. We consider an urn with balls of five colors: white, black, red, yellow, green. A hundred draws with replacement have given the following counts: (20, 5, 10, 40, 25). A ball will be drawn from the urn and you are asked to predict its color. Your gain will depend on your choice and of the true color, according to the following payoff matrix:

		True color				
		W	B	R	Y	G
Choice	W	100	50	30	30	20
	B	300	500	200	100	-100
	R	50	100	200	100	50
	Y	0	10	20	50	20
	G	-20	-50	100	150	200

Compute the lower, upper and pignistic utilities for each choice. Which color do you choose, using different decision rules? (Hint: write a function that computes the lower, upper and pignistic expected utilities given a mass function and a utility matrix).