PMPC Tutorial Sheet 8

1. Data are collected in a paired-comparison experiment. On each trial subjects have to say which of two options they prefer. Assume that the utility of each option i in a set of n options varies randomly on each trial and that the subject picks the options with the higher utility. The utility for each option has a Normal distribution with mean μ_i and standard deviation 1. Let $d_{ij} := \mu_i - \mu_j$ be the difference in utility between the mean utilities of two options. The probability of choosing i over j is then[1]

$$p_{ij} = 1 - \Phi(0; d_{ij}, \sqrt{2}) = \Phi(d_{ij}/\sqrt{2}; 0, 1).$$

   Let there be three options with mean utilities $\mu_1 = 1$, $\mu_2 = 1.5$, $\mu_3 = 2.5$. Calculate the choice probabilities for all possible pairs of options and put them into a matrix. How often will it happen that a subject prefers 2 over 1 and 3 over 2 but also 1 over 3?

2. Rumelhart and Greeno [2] conducted a paired-comparison experiment to test various choice models. They considered all pairs of 9 celebrities: Lyndon Johnson, Harold Wilson, Charles DeGaulle, Johnny Unitas, Carl Yastrzemski, A. J. Foyt, Brigitte Bardot, Elizabeth Taylor, Sophia Loren. The first three were politicians, the second three athletes, and the remaining three actresses. They presented each of 234 subjects with hypothetical choices between all possible pairings of the nine, asking for each pair with whom they would rather spend an hour of conversation with. The data are shown in table 1 (that you can also download from Stud.IP). Rows were chosen over columns. Hence, the first row shows how often Johnson was chosen over Wilson, over DeGaulle, and so on. A celebrity was, of course, never paired with itself and data on the diagonal

[1] The cumulative normal distribution Φ is called `normcdf` in Matlab and Octave and it's inverse is `norminv`.

	LJ	HW	CD	JH	CY	AF	BB	ET	SL
LJ	117	159	163	175	183	179	173	160	142
HW	75	117	138	164	172	160	156	122	122
CD	71	96	117	145	157	138	140	122	120
JH	59	70	89	117	176	115	124	86	61
CY	51	62	77	58	117	77	95	72	61
AF	55	74	96	119	157	117	134	92	71
BB	61	78	94	110	139	100	117	67	48
ET	74	112	112	148	162	142	167	117	87
SL	92	112	114	173	173	163	186	147	117

Table 1: Data from a paired comparison experiment [2]. The data show how often the rows were preferred over the columns for 234 subjects that did all pairwise comparisons each. The diagonal was simply filled up with $\frac{234}{2}$ for convenience.

of this matrix were actually not collected. I filled in the theoretical value of $\frac{1}{2} \cdot 234$ to simplify the analysis. Apply Thurstonian scaling to these data. Who is the most popular celebrity in this set, who is the least popular? Make a plot of the observed relative frequency of the choices as a function of the fitted probabilities.

3. The triple HW (2), SL (9), ET (8) violates strong stochastic transitivity. $p_{29} \geq \frac{1}{2}$ and $p_{98} \geq \frac{1}{2}$ but $p_{28} < p_{98}$. Explain what this means in your own words. Identify all triples in the data that violate strong stochastic transitivity. Why are there violations of strong stochastic transitivity in these data?

4. Bonus question: Write a function `restle.m` in Matlab or Octave that implements Restle's choice model [1]. The function gets the log of the utilities for each feature (the log is important so that the inputs can be positive and negative), a feature matrix, and a choice matrix (like the one in table 1) as inputs and returns the negative log likelihood and the theoretical choice probabilities as outputs.[2] What is an appropriate feature matrix for the data in table 1? Use the function `fminunc` to find the values for the feature utilities that minimize the negative log likelihood. For a function `f` of an n-dimensional vector `x` the following call finds the `x` that minimizes `f`: `x = fminunc(f, zeros(1,n))`[3]. In this example the function `fminunc` starts a numerical search for the best value with an initial value of all zeros. In order to define an appropriate `f` as an anonymous function that can be used to minimize the negative log likelihood the following line will be useful: `f = @(logutilities) restle(logutilities, featureMatrix, data)`. `f` is now a variable that refers to a function that takes `logutilites` as input but fixes the feature matrix and the data. You can use such an `f` as input to `fminunc`. Make a plot of the observed relative frequency of the choices as a function of the fitted probabilities.

References

[1] F. Restle. *Psychology of Judgment and Choice: A Theoretical Essay.* John Wiley & Sons, 1961.

[2] D. L. Rumelhart and J. G. Greeno. Similarity between stimuli: An experimental test of the luce and restle choice models. *Journal of Mathematical Psychology*, 8:370–381, 1971.

[2] Or download the function `restle.m` from Stud.IP; If you do implement it yourself, make sure that the diagonal of the choice matrix contains not `NaN` but $\frac{1}{2}$.

[3] `fminunc` is for Octave or Matlab if you have the optimization toolbox. Normally, in Matlab you will have to use `fminsearch` instead. Use the following code: `o = optimset; o.MaxFunEvals = 100000; o.MaxIter = 100000; o.TolX = 10^-8; o.TolFun = 10^-8; x = fminsearch(f,zeros(1,n),o);`