Optimized Democracy (Spring 2021)
Assignment #2

Ariel Procaccia (instructor) and Jamie Tucker-Foltz (TF)

Due: 3/2/2021 11:59pm ET

Instructions:

- It is fine to look up a complicated sum or inequality, but please do not look up an entire solution. In particular, the solutions to many of the problems that we give can be found in papers, but, needless to say, you should avoid reading the proof if you come across the relevant paper. If for some reason you did see the solution before working it out yourself, please say so in your solution.
- You may discuss the problems with classmates but please write down solutions completely on your own.
- Please *type up* your solution and send a PDF to Jamie by email (`jtuckerfoltz@gmail.com`), including "CS238HW2" somewhere in the subject. Attach your solutions as a PDF of the form "FirstLast.pdf", where "First" is replaced by your preferred first name and "Last" is replaced by your last name.

Problems:

1. In class we discussed the Mallows model, which gives an expression for the probability of a ranking σ given the ground truth π. So computing the probability of a given ranking is easy, but how can we sample from this distribution?

   Assume that $a_1 \succ_\pi a_2 \succ_\pi \cdots \succ_\pi a_m$, and consider the following generative model, defined by probabilities p_{ij} for all $i = 1, \ldots, m$ and $j = 1, \ldots, i$, which iteratively constructs the ranking σ. In round 1, a_1 is inserted into the first (and only) position of the constructed ranking with probability $p_{11} = 1$. In round 2, a_2 is inserted into position 1 (above a_1) with probability p_{21} and into position 2 (below a_1) with probability p_{22}. More generally, in round i, for each $j = 1, \ldots, i$, a_i is inserted into position j with probability p_{ij}.

   [30 points] Prove that the Mallows Model with parameter ϕ is equivalent to this generative model with $p_{ij} = \phi^{i-j} \frac{1-\phi}{1-\phi^i}$. (This means that sampling rankings from the Mallows model is indeed easy.)

   Hint: You may use the fact that for all $\pi \in \mathcal{L}$,

   $$(1+\phi)(1+\phi+\phi^2)\cdots(1+\phi+\cdots+\phi^{m-1}) = \sum_{\tau \in \mathcal{L}} \phi^{d_{KT}(\tau,\pi)}.$$

2. A shortcoming of the epistemic approach we discussed in class is that the "optimal" rule depends on the details of the noise model. For example, Kemeny is an MLE with respect to the Mallows model, but wouldn't be an MLE if the noise had a different form. In this problem we will instead explore a worst-case epistemic approach.

   Let $\mathcal{L}$ be the set of rankings over alternatives. Let

$$d : \mathcal{L} \times \mathcal{L} \to [0, \infty)$$

   be a *metric* over $\mathcal{L}$, which means that, for any rankings $\sigma_1, \sigma_2, \sigma_3 \in \mathcal{L}$,

   - $d(\sigma_1, \sigma_2) = 0 \iff \sigma_1 = \sigma_2$,
   - $d(\sigma_1, \sigma_2) = d(\sigma_2, \sigma_1)$, and
   - $d(\sigma_1, \sigma_3) \leq d(\sigma_1, \sigma_2) + d(\sigma_2, \sigma_3)$ (this is called the *triangle inequality*).

   Think of d as an abstract way of measuring the distance between two preference rankings. For example, it is easily verified that d_{KT} satisfies all three axioms.

   Suppose that there is some ground truth ranking $\pi \in \mathcal{L}$, and we are given an input preference profile $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathcal{L}^n$ with the guarantee that the average distance between π and rankings in $\boldsymbol{\sigma}$ is at most some constant $t \geq 0$. That is, we assume

$$\pi \in B_t(\boldsymbol{\sigma}) := \{\tau \in \mathcal{L} \mid d(\boldsymbol{\sigma}, \tau) \leq t\},$$

   where

$$d(\boldsymbol{\sigma}, \tau) := \frac{1}{n} \sum_{i \in N} d(\sigma_i, \tau).$$

   We then choose $\widehat{\pi} \in \mathcal{L}$ to minimize the worst-case distance from $\widehat{\pi}$ to the unknown ground truth π. Under this process, the worst-case distance to the ground truth is given by

$$k := \max_{\boldsymbol{\sigma} \in \mathcal{L}^n} \min_{\widehat{\pi} \in \mathcal{L}} \max_{\pi \in B_t(\boldsymbol{\sigma})} d(\widehat{\pi}, \pi).$$

   We seek to understand the behavior of k as a function of t (as it turns out, n and d don't matter too much).

   Prove the following statements:

   (a) **[8 points]** $k \leq 2t$. In words, given $\boldsymbol{\sigma}$ whose average distance from π is at most t it is always possible to find a ranking $\widehat{\pi} \in \mathcal{L}$ that is guaranteed to be at distance at most $2t$ from π.

   (b) **[12 points]** Suppose that, instead of allowing for an arbitrary $\widehat{\pi} \in \mathcal{L}$, we require that $\widehat{\pi}$ be one of the rankings σ_i of the input profile $\boldsymbol{\sigma}$, i.e., define

$$k' := \max_{\boldsymbol{\sigma} \in \mathcal{L}^n} \min_{i \in N} \max_{\pi \in B_t(\boldsymbol{\sigma})} d(\sigma_i, \pi).$$

   Then $k' \leq 3t$.

   (c) **[8 points]** Assume that t is in the image of d. Then $k \geq \frac{t}{2}$.

(d) **[12 points]** Assume that t is in the image of d, and that d is neutral (i.e., the distance between two rankings is invariant to renaming the alternatives). Then $k \geq t$.

3. In this question, we analyze how an organization can use voting to decide how to divide its budget among projects. For example, a student organization which collects membership dues might poll its members which types of events they are interested in attending, and use those votes to allocate funding to the event types.

   Let A be a set of alternatives (or projects). An approval profile $P : N \to 2^A$ is a function that assigns to each voter $i \in N$ a non-empty approval set. An approval-based *budget division rule* f is a function that takes as input an approval profile P and returns a distribution $p : A \to [0, 1]$ with $\sum_{x \in A} p_x = 1$. We interpret p_x as the fraction of the budget spent on x. The utility of voter i obtained from p is $u_i(p) = \sum_{x \in P(i)} p_x$, the fraction spent on approved projects.

   (a) **[5 points]** Consider the *utilitarian rule* f_{util} which returns the uniform distribution over the approval winners of a profile. Since approval voting is strategyproof, it is easy to see that f_{util} is also strategyproof.

      i. Prove that f_{util} is *Pareto optimal*: for every P, there does not exist a distribution q such that $u_i(q) \geq u_i(f_{\text{util}}(P))$ for all $i \in N$, and $u_i(q) > u_i(f_{\text{util}}(P))$ for some $i \in N$.

      ii. Show that f_{util} can be unfair to voters: Give an example of a profile P such that $u_i(f_{\text{util}}(P)) = 0$ for some voter $i \in N$.

   (b) **[15 points]** Let f be a budget division rule that is Pareto optimal and *minimally fair*: for all profiles P, we have $u_i(f(P)) > 0$ for all $i \in N$. Consider the following two profiles:

   $$P = (\{a\}, \{a, c\}, \{a, d\}, \{b, c\}, \{a, b\}),$$
   $$P' = (\{a\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}),$$

   and write $p = f(P)$ and $p' = f(P')$. Note that in P, alternatives b and c are symmetric, since if we swap their names we get an identical profile up to reordering voters. Similarly, in P', alternatives c and d are symmetric. Let us assume that f is a *symmetric* rule, which for this case means $p_b = p_c$ and $p'_c = p'_d$.

   Show that $p'_a + p'_b > p_a + p_b$. Deduce that every symmetric, Pareto optimal, and minimally fair rule can be manipulated.

   (c) **[10 points]** Show that there exists a strategyproof rule that is minimally fair, and show that there exists a Pareto optimal rule that is minimally fair.