ST790 – Homework 1 Due: 11/22/2022

These exercises are meant to supplement the lectures by providing some further examples to illustrate the general ideas and theory. Students (individually or in pairs) should attempt to solve all the assigned problems. The solutions will be collected at the end of the semester (the day before Thanksgiving), so you may work on these at your own pace. But don't wait too long to get started! If you have questions, feel free to ask.

1. Justify the claim made in the Week 01a lecture (slide 19) that the most diffuse probability distribution in the credal set $\mathscr{C}(\overline{\Pi}_x)$ is $\Pi_x^\star = \mathsf{N}(x, 1)$. Your justification doesn't have to be a formal proof, just reason it out from the fact that $\Pi_x \in \mathscr{C}(\overline{\Pi}_x)$ if and only if $\Pi_x(A) \leq \overline{\Pi}_x(A)$ for all $A \subseteq \mathbb{R}$.

   Hint: For sets A of the form $A = (-\infty, x-a] \cup [x+a, \infty)$, with $a > 0$, the equality is attained, but for any other sets A, it's a strict inequality.

2. For a fairly general class of continuous, scalar parameter problems, there is a simple formula for Fisher's fiducial distribution.[1] Suppose the model for data $X \in \mathbb{R}$ has an absolutely continuous distribution function F_θ, for $\theta \in \mathbb{R}$. Then the fiducial distribution Ψ_x for θ, given $X = x$, has a density function given by

$$\psi_x(\theta) = \Bigl|\frac{\partial F_\theta(x)}{\partial \theta}\Bigr|, \quad \theta \in \mathbb{R}.$$

   (a) Suppose that X is an exponential random variable with rate parameter $\theta > 0$, so that $F_\theta(x) = 1 - e^{-\theta x}$, for $x > 0$. Use the formula above to find the fiducial distribution Ψ_x of θ, given $X = x$. Verify that $\psi_x(\theta)$ is, indeed, a density function in θ for fixed x.

   (b) Define the p-value function $\pi_x(\theta) = \mathsf{P}_\theta\{|\log(\theta X)| \geq |\log(\theta x)|\}$. Draw a plot of this function of θ when $x = 0.25$.

   Hint: You can simplify the p-value by setting $g(z) = \mathsf{P}_\theta\{|\log(\theta X)| > z\}$ first, which doesn't depend on θ (why?), and then writing $\pi_x(\theta) = g(|\log(\theta x)|)$.

   (c) As in the Week 01a lecture, define the corresponding upper probability as

$$\overline{\Pi}_x(A) = \sup_{\theta \in A} \pi_x(\theta), \quad A \subseteq [0, \infty).$$

   Argue, like in the previous problem (numerical justification is fine), that the fiducial distribution Ψ_x is in the credal set $\mathscr{C}(\overline{\Pi}_x)$ determined by $\overline{\Pi}_x$.
   Hint: Equality $\Psi_x(A) = \overline{\Pi}_x(A)$ is attained for sets A of the form $A = [0, a/x] \cup [1/ax, \infty)$, for $a \in [0, 1]$, but strict inequality for all other A.

3. (From Section 3.3 in Nguyen's *An Introduction to Random Sets*, 2006.) Consider an box that 30 red balls and 60 other balls, some of which are white and the others are black. Let X denote the color of a ball chosen at random from this box. Obviously, this isn't enough information to determine the distribution of X; all you can say is that X has one of the mass functions f_k, for $k = 0, 1, \ldots, 60$, where

[1]See Section 3 in Zabell's "R.A. Fisher and the fiducial argument," *Statistical Science*, 1992.

x	red	black	white
$f_k(x)$	$30/90$	$k/90$	$(60-k)/90$

(a) Say a ball is "dark" if it's not white. What's the lower and upper probability of drawing a dark ball?

(b) Sketch the probability simplex as displayed in the Week 01a lecture and, in it, the corresponding set of precise probabilities. Is this a credal set?

(c) Suppose that there's a payoff depending on the color of the ball that's drawn. In particular, the player wins \$30 for a red ball, \$20 for a black ball, and \$15 for a white ball. Find the lower and upper expected winnings, and explain your rationale.

4. Here's the standard example of a probability that's finitely additive but not countably additive. Consider the universe Ω consisting of the natural numbers (or any other countably infinite set) and, for any subset $A \subseteq \Omega$, define the probability

$$\mathsf{P}(A) = \begin{cases} 0 & \text{if } A \text{ is finite} \\ 1 & \text{if } A \text{ is co-finite (complement of a finite set).} \end{cases}$$

(a) Show that P is finitely additive in the sense that

$$\mathsf{P}\Big(\bigcup_{k=1}^{K} A_k\Big) = \sum_{k=1}^{K} \mathsf{P}(A_k) \quad \text{for all } K < \infty \text{ and all disjoint } A_1, \ldots, A_K,$$

but not countably additive, i.e.,

$$\mathsf{P}\Big(\bigcup_{k=1}^{\infty} A_k\Big) \neq \sum_{k=1}^{\infty} \mathsf{P}(A_k), \quad \text{for all disjoint } A_1, A_2, \ldots.$$

(b) Argue that P does not have a mass function. That is, argue that there is no function p such that $\mathsf{P}(A) = \sum_{\omega \in A} p(\omega)$ for all $A \subseteq \Omega$.

5. Let γ be a capacity and $\tilde{\gamma}$ its dual/conjugate. Consider the claim

If γ is super-additive, then $\tilde{\gamma}$ is sub-additive.

Show, by constructing a counter-example, that this claim is false.

6. On the three-state universe $\{x_1, x_2, x_3\}$, define the lower probability $\underline{\mathsf{P}}$

$$\underline{\mathsf{P}}(x_1) = 0.45, \quad \underline{\mathsf{P}}(x_2) = 0.15, \quad \underline{\mathsf{P}}(x_3) = 0.30,$$

and corresponding upper probability $\overline{\mathsf{P}}$

$$\overline{\mathsf{P}}(x_1) = 0.55, \quad \overline{\mathsf{P}}(x_2) = 0.20, \quad \overline{\mathsf{P}}(x_3) = 0.40.$$

(a) Recall that $\underline{\mathsf{P}}$ is 2-monotone if

$$\underline{\mathsf{P}}(A \cup B) + \underline{\mathsf{P}}(A \cap B) \geq \underline{\mathsf{P}}(A) + \underline{\mathsf{P}}(B), \quad \text{for all } A, B.$$

Check the above inequality for as many pairs of subsets (A, B) as it takes to convince yourself that $\underline{\mathsf{P}}$ is 2-monotone.

Hints: First, recall that $(\underline{\mathsf{P}}, \overline{\mathsf{P}})$ are dual; second, while there are totally 64 pairs of subsets, but you don't have to check them all: the inequality is an obvious equality if $A = B$ and, moreover, it holds for (A, B) if and only if it holds for (B, A), so there is some structure.

(b) Sketch the probability simplex $\mathscr{P}$ like in the Week 01a lecture and, on it, plot the credal set $\mathscr{P}_0 \subset \mathscr{P}$ determined by $(\underline{\mathsf{P}}, \overline{\mathsf{P}})$. That is, sketch the region

$$\mathscr{P}_0 = \{(p_1, p_2, p_3) \in \mathscr{P} : p_1 \in [0.45, 0.55], p_2 \in [0.15, 0.20], p_3 \in [0.30, 0.40]\}.$$

Hint: It may help to sketch the simplex in *Barycentric coordinates.*[2]

(c) It's clear from the sketch in Part (b) that $\mathscr{P}_0$ is non-empty. But if the space was more complex, then drawing the sketch wouldn't be possible. To know $\mathscr{P}_0$, it's enough to find the vertices—there are six of them in this case. Fortunately, there is a general way to identify the extreme points of the credal set determined by a 2-monotone capacity.

Suppose the space is $\{x_1, \ldots, x_K\}$, and let σ denote a permutation of the indices $\{1, \ldots, K\}$. For a given σ, define the probability vector

$$p_k^{(\sigma)} = \underline{\mathsf{P}}(\{x_{\sigma(1)}, \ldots, x_{\sigma(k)}\}) - \underline{\mathsf{P}}(\{x_{\sigma(1)}, \ldots, x_{\sigma(k-1)}\}), \quad k = 1, \ldots, K.$$

Then Chateauneuf & Jaffray, *Mathematical Social Sciences*, 1989, showed that the vectors $\{p^{(\sigma)} : \text{all permutations } \sigma\}$ is the set of extreme points. Pick any one of the $3! = 6$ permutations σ, find the corresponding $p^{(\sigma)}$ vector, and identify the the corresponding vertex of the credal set you sketched in Part (b) above; you can do more than one σ if you like.

7. In the Week 01a lecture, from a p-value function $\pi_x(\vartheta)$ for testing a point null hypothesis $H_0 : \theta = \vartheta$, a p-value function for general hypotheses A was defined as

$$\overline{\Pi}_x(A) = \sup_{\vartheta \in A} \pi_x(\vartheta).$$

If it helps, you can focus specifically on the scalar normal mean case.

(a) Show that $\overline{\Pi}_x$ is monotone and sub-additive.

(b) Show that $\overline{\Pi}_x$ is 2-alternating or, equivalently, that its dual $\underline{\Pi}_x$ is 2-monotone.

(c) P-values are often criticized for (allegedly) lacking the mathematical niceties that probabilities possess. Parts (a) and (b) above, however, showed that p-values actually do have mathematical structure,[3] it's just different from probabilities. So there must be some confusion here somewhere...

[2] e.g., `https://en.wikipedia.org/wiki/Barycentric_coordinate_system`

[3] In fact, p-values as defined above are possibility measures, so they have even nicer properties than 2-alternating capacities.

One such criticism is in Mark Schervish's "P-values: What they are and what they are not," *The American Statistician*, 1996.[4] He argues that p-values are "incoherent" and not "measures of support" by constructing an example where monotonicity of the p-value function is violated. This is incompatible with the conclusion of Part (a) above, so something weird is going on. Can you spot how Schervish's p-value construction differs from that above?

[4]e.g., `https://www.apps.stat.vt.edu/leman/VTCourses/schervish-pvals.pdf`