Details from the Week 03a example*

Ryan Martin

09/08/2022

Recall that the setup is as follows. Consider a ε-contamination neighborhood

$$\mathscr{P} = \{\mathsf{P} = (1-\varepsilon)\mathsf{P}_0 + \varepsilon\mathsf{Q} : \text{any probability } \mathsf{Q}\},$$

where $\varepsilon \in (0,1)$ and P_0 a probability distribution on $\mathbb{X}$ are given. This is a common model used in the robust statistics literature. Define the lower envelope of $\mathscr{P}$ as

$$\underline{\Pi}(A) = \inf_{\mathsf{P} \in \mathscr{P}} \mathsf{P}(A), \quad \text{for all measurable } A \subseteq \mathbb{X}.$$

Define the credal set $\mathscr{C}(\underline{\Pi}) := \{\mathsf{P} : \mathsf{P}(\cdot) \geq \underline{\Pi}(\cdot)\}$ corresponding to the lower envelope, i.e., all probabilities lower-bounded by $\underline{\Pi}$. Then the goal is to prove the following

Claim. $\mathscr{P} = \mathscr{C}(\underline{\Pi})$.

Proof. The proof shows equality of the two sets of probabilities by showing (a) $\mathscr{P} \subseteq \mathscr{C}(\underline{\Pi})$ and (b) $\mathscr{P} \supseteq \mathscr{C}(\underline{\Pi})$. For Part (a), start by taking $\mathsf{P} \in \mathscr{P}$. Since $\underline{\Pi}$ is defined as the event-wise infimum, there is no A such that $\mathsf{P}(A)$ is less than $\underline{\Pi}(A)$. Evidently, P is lower-bounded by $\underline{\Pi}$, so we conclude that $\mathsf{P} \in \mathscr{C}(\underline{\Pi})$.[1]

For Part (b), first note that

$$\underline{\Pi}(A) = \inf_{\mathsf{P} \in \mathscr{P}} \mathsf{P}(A) = \inf_{\mathsf{Q}} \{(1-\varepsilon)\,\mathsf{P}_0(A) + \varepsilon\,\mathsf{Q}(A)\} = (1-\varepsilon)\,\mathsf{P}_0(A).$$

So if $\mathsf{P} \in \mathscr{C}(\underline{\Pi})$, then we have that

$$\mathsf{P}(A) \geq (1-\varepsilon)\,\mathsf{P}_0(A), \quad \text{for all } A. \tag{1}$$

From this it's easy to check that

$$\mathsf{Q}_\varepsilon(A) := \mathsf{P}(A) - (1-\varepsilon)\,\mathsf{P}_0(A), \quad A \subseteq \mathbb{X},$$

is a finite measure, i.e., $\mathsf{Q}_\varepsilon(\varnothing) = 0$, $\mathsf{Q}_\varepsilon(\cdot) \geq 0$, $\mathsf{Q}_\varepsilon(\mathbb{X}) = \varepsilon$, and it's countably additive.[2] I claim that Q_ε can be rescaled to be a probability measure, i.e., that $\varepsilon^{-1}\mathsf{Q}_\varepsilon$ is a probability measure. If this is true, then I can take $\mathsf{Q} = \varepsilon^{-1}\mathsf{Q}_\varepsilon$ and write

$$\mathsf{P}(A) = (1-\varepsilon)\,\mathsf{P}_0(A) + \varepsilon\,\mathsf{Q}(A),$$

*Sorry for being careless with this example. The argument I gave in lecture is more-or-less fine, I just overlooked a detail that prevented me from being able to answer the question that Shubhajit raised.

[1] Note that this part of the proof didn't rely on the form of $\mathscr{P}$ or $\underline{\Pi}$, so the containment $\mathscr{P} \subseteq \mathscr{C}(\underline{\Pi})$ always holds, simply by definition of the lower envelope.

[2] A linear combination of countably additive things is countably additive.

which will prove $\mathsf{P} \in \mathscr{P}$. To check the claim that $\mathsf{Q}_\varepsilon(\cdot) \leq \varepsilon$, consider a situation like Shubhajit suggested in lecture, e.g., where A is such that $\mathsf{P}_0(A) = 0$ and $\mathsf{P}(A) = \delta > \varepsilon$. In this case, we'd have

$$\mathsf{P}(A^c) - (1 - \varepsilon)\,\mathsf{P}_0(A^c) = (1 - \delta) - (1 - \varepsilon) = \varepsilon - \delta < 0,$$

which contradicts (1). So it must be that $\mathsf{Q}_\varepsilon \leq \varepsilon$ and, therefore, rescaling by dividing by ε turns it into a probability measure, from which we conclude $\mathsf{P} \in \mathscr{P}$. $\square$