Evaluating infectious disease forecasts with allocation scoring rules

2023-03-28

Introduction

Infectious disease forecasts have been used as an input to public health decision-making processes such as …

In this manuscript we argue that for many decision-making problems, it is possible and valuable to develop scoring rules for probabilistic forecasts that are responsive to that decision-making context.

Literature review

Previous work in infectious disease forecasting where applications of forecasting to decision-making problems have been developed

Theoretic work on decision-theoretic set up for forecast evaluation

Previous applications of decision-theoretic evaluation to fields like economics

Previous decision-theoretic work in infectious disease forecast evaluation. I think there are at least a couple of papers out there at some level of formality. Maybe Kat Shea has something?

operations research work on constrained allocation

Methods

A review of probabilistic forecasting and the decision-theoretic set up for scoring rules

General notation. scoring functions, scoring rules.

A high-level/intuitive discussion of the three-step recipe for obtaining scoring rules from loss/utility in a decision-making context.

1. Specify down a scoring function that measures the loss associated with taking action x when outcome y materializes

2. Given a probabilistic forecast F, determine the Bayes act

3. The scoring rule calculates the score as the loss incurred when the Bayes act was used

A review of quantile loss, CRPS, and the weighted interval score

Set up of decision-making problem with costs for over-/under-prediction. Quantile/pinball loss.

Statement that this loss elicits quantiles as the Bayes act, derivations deferred to a supplement.

Interpretation of CRPS as an integral of quantile loss over quantile levels, WIS as a discrete approximation of CRPS.

Allocation scoring rules

Set of decision-making problem with costs for over-/under-prediction across multiple locations and a constraint on the total allocation. Allocation loss.

Description of the Bayes act at some level (derivations go in a supplement):

- divide available resources across locations so that according to F, the expected benefit of 1 additional unit is the same everywhere.

- Statement of the result if $g_i(x) = x$.

A figure with one or two sub-panels to illustrate. What figures?

Description of what happens if there is uncertainty about the costs O, U, or the constraint K. Integration across those results in something like a forecaster-specific weighted CRPS.

Applications

Application with a quantile elicitation problem

Consider one of the following problems?

- "should we expand hospital bed capacity"

- "should impose a stay at home order".

for individual locations with specific costs for taking the action and possible costs for "under-acting". Both are I think hard to operationalize in the sense that it's not immediate what specifically those costs should be.

Application with an allocation problem

Hospital admissions in the U.S.

Case study heading into the Omicron wave. Some more detailed discussion of implications of bad forecasts for specific decision-making purposes – take a "deep dive" into one or two example states like FL.

Look at results over a broader range of time.

Discussion

We often conceive of infectious disease forecasts as being useful for decision-making purposes, but it is rare for forecast evaluation to be tied directly to the value of the forecasts for informing those decisions. This work seeks to address that gap.

We have demonstrated that evaluation methods that are tied to decision-making context can yield model rankings that are substantively different from generic measures of forecast skill like WIS.

In practice, there are many users of forecasts with many different decision-making problems. Not all can be easily quantified. Those that can be easily quantified may differ enough that no single score is appropriate for all users. We suggest reporting multiple scores. This may be tricky to operationalize in the setting of a general forecast hub. It matters how you elicit and represent probabilistic forecasts (quantiles? samples? cdfs?).

The allocation score we developed here does not directly account for important considerations such as fairness/equity of allocations.

The allocation score we developed also does not attempt to capture the broader context of decision-making. For example, in practice it may be possible to increase the resource constraint K by shifting funding from other disease mitigation measures.

There is much more to do in this general area.

References