Binary to continuous methodology draft

2023-04-29

We begin with a basic task in public health: triage at the unit level, that is, the problem of deciding to whom, where, and when a unit of scarce health-care resource is allocated. We assume that the resource unit, e.g., a bed, ventillator, dose, or clinician, is indivisible, and that a potential recipient is a single person during a single period of time (say, a day). We also focus on the scenario of a recipient that is not yet under care, such as a person that may be infected with SARS-COV-2 in the future and consequently require hospitalization or intubation.

A decision problem is…

The unit-level decision problem:

Assume that there are currently resources allocated to care for x_a and x_b many cases in locations a and b. Another resource unit becomes availible and we must decide whether to allocate it to a new potential case in location a or b. In terms of a decision variable d the decision is whether to set $d = a$ or $d = b$.

To support our decision making process we elicit forecasts p_a of $P(Y_a = 1)$ and p_b of $P(Y_b = 1)$ where $Y_a, Y_b = 1$ if a case in addition to what is currently provided for at horizon h occurs during the day at horizon h in a, b, and 0 otherwise.

$p_{a,b}$ can in particular come from a distributional incidence forecast F_a via $p = 1 - F_a(x)$ where x is the number of cases currently provided for in h days.

- fix prob β of additional case in b, via
  - baseline/persistence/climatology, or
  - F_b taken as a given, not to be evaluated, or
  - $F_{\sum i \neq a}(x) = \int_{\mathbb{R}^n} \mathbb{1}\{\sum_{i \neq a} x_i \leq x\} dF$ if b is a collection of locations and F is a given joint distribution for all coations other than a.

We define a loss function for the decision problem that encodes the dilemma faced when balancing the risk of a case of unmet need in a incurred when $d = b$ is chosen against the guaranteed "average" unmet need of β in b that is incurred when $d = a$ is chosen. This is the function of the decision d and outcome Y_a

$$l(d, Y) = \mathbb{1}\{d = a\}\beta + \mathbb{1}\{d = b\}Y_a,$$

which picks one of the 4 entries in the table

	New case in a:	Yes	No
allocate to:	a	β	β
	b	1	0

Associated with this loss function is the *regret* function

$$r(d, Y) = \mathbb{1}\{d = a\}\beta + (\mathbb{1}\{d = b\} - \beta)Y_a,$$

with tabular form

	New case in a:	Yes	No
allocate to:	a	0	β
	b	$1 - \beta$	0

This encodes the allocation dilemma in terms of how much worse our decision is than that of an oracle that only allocates to a when $Y_a = 1$.

According to the forecast p, the decision has expected loss and regret

$$E_p[l(d)] = \mathbb{1}\{d = a\}\beta + \mathbb{1}\{d = b\}p \tag{1}$$
$$E_p[r(d)] = \mathbb{1}\{d = a\}\beta + (\mathbb{1}\{d = b\} - \beta)p \tag{2}$$

which are minimized by the decision rule

$$d(p) = \begin{cases} a & \text{if } p > \beta \\ b & \text{otherwise.} \end{cases}$$

This decision — or *Bayes* — rule allows us to convert l and r into a scoring functions

$$s_l(p, Y_a) = l(d(p), Y_a) = \mathbb{1}\{p > \beta\}\beta + \mathbb{1}\{p \leq \beta\}Y_a \tag{3}$$
$$s_r(p, Y_a) = s_l(p, Y_a) - \beta Y_a \tag{4}$$

- Show, a la Ehm eqs 13-16 that integrating over x gives a Brier score for any β
  - relate to quantile score for F_a
- Show how looking at different β with x fixed gives a *value* or *Murphy* curve
  - integrating over x now gives CRPS...