Allocation scores, WIS, and CRPS via decision theory

Aaron Gerding, Nicholas G. Reich, Benjamin Rogers, Evan L. Ray

August 18, 2023

Abstract

A somewhat open-ended attempt to motivate and construct quantile scores, WIS, CRPS, and allocation scores within a unified decision theoretic framwork.

1 Introduction

In this section/report/paper/rambling manifesto, we present a methodological perspective on the use scoring rules for forecast evaluation that we see as well-adapted to our motivating research goal to understand and describe the social utility of infectious disease forecasts. Our approach follows the common practice in decision theory of giving central roles to the foundational concepts of state, action, and consequence. From these primitives, we first derive the familiar measures of quantile score (QS), continuous ranked probability score (CRPS) and their aggregates across forecast targets. We then introduce new scoring rules, the *specific* and *integrated allocation scores*, as reparametrizations of QS and CRPS aggregates that result when part of our definition of consequence is shifted into our definition of action. Specifically, we will trade the regret a decision maker experiences as a consequence of seeing allocated resources go unused for the requirement that a decision maker choose an action that allocates no more than a certain amount in total. This exchange mirrors the duality in mathematical optimization theory between constraints on "decision variables" and the "Lagrange multiplier variables" which quantify the sensitivity of solutions to the constraints, that is, the regret (or relief) experienced when a change in the constraint leads to a change in the optimal cost. Our computational methods (see ..) in fact make use of this connection. Something about Bayes risks...

2 Scoring forecasts via decision problems

Broadly speaking, we score a probabilistic forecast F by identifying it with a solution to a *decision problem* made up of pre-specified actions, states and consequences and rating the utility of that solution. That is, F will receive a "better" or "worse" score in a particular instance when the consequence is better or worse of a rational decision maker using F (and essentially only F) to select an action in the face of uncertainty about a future state. It should be noted immediately that a better or worse score in a single instance may tell us almost nothing about the intrinsic value of the forecasting method that produced F. But as scores accumulate, we can use their distributions across forecasting instances to make inferences as to whether one forecasting method is better or worse than another in a decision theoretic sense.

2.1 The ingredients of a decision problem with examples

The basic example of a future state we are considering is the number of individuals that become sick enough from an infectious disease to be in need of some potentially limited resource such as hospital beds, ventilators, medication, or medical staff. We write y to refer to a particular state such as the quantity of need in a single location or the quantities across multiple locations or times and Y for a future uncertain state which is a random variable with a set of realizable values $\mathcal{Y}$. The forecast F to be scored is a claim of how Y is (or will be) distributed. We also will refer to states as outcomes, which helps to emphasize their future realization. But we want to clearly distinguish an outcome from a consequence, which as we discuss below, is the result of taking an action under a certain outcome.

The basic example of an action we consider is the specific amount or set of amounts of such a resource that a public health worker decides to make available in either a single location or across multiple locations or times. We write an action as x and assume that it takes values in a set of possible actions $\mathcal{X}$ that is possibly defined by some constraint such as the total amount of resource available to the public health worker for allocation. Note that x must be selected prior to the realization of y. A decision maker might use a set rule to select x based on inputs, such as F, but the observed value of y cannot be such an input.[1]

And when actions are assigned to states, there are consequences, which we abstractly identify with an element c of a set $\mathcal{C}$. We have two basic examples in mind of what such a c might be. First, we consider some amount of the regret mentioned above that resources made available in a place or time are not actually needed at that place or time. We emphasize that the nature of this regret is potentially quite amorphous. In business contexts (e.g., the "newsvendor problem" from inventory management) this regret is often easily quantifiable as monetary waste, but in a public health context, it could... (A desire to better understand over-allocation regret and its relevance to infectious disease forecast evaluation was actually a primary impetus for the present work.) Our second and more concrete example of a consequence is the loss incurred when a sick individual is not able to access a resource because it was not made available to them where and when they needed it. We emphasize here that while we formally identify such loss as a consequence of a particular action being taken under a particular state, it may or may not be an avoidable loss. That is, it could just as well result from a factor outside a decision maker's control such as a severe resource constraint as from a choice to withhold a resource, motivated, for example, by a desire to reduce waste.

2.2 The formal procedure

With a view towards a decision problem defined by these examples of sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{C}$ of actions, outcomes, and consequences, we now give a formal three-step procedure for defining and evaluating a *scoring rule* for a forecast F for a future value $y \in \mathcal{Y}$ of Y.

1. Specify a *loss function* $s(x, y)$ that quantifies the utility or disutility of the consequence c of taking action x under outcome y.

2. Map the given probabilistic forecast F for Y to an action x^F which "solves" the decision problem by minimizing the expected loss $E_F[s(x, Y)]$ over all $x \in \mathcal{X}$ under the distribution F. Following **?**, we call x^F the *Bayes act* for F. Note that x^F depends not only on F and s, but also on the

---

[1]In a dynamical decision theoretic approach, the action x could be a time series depending at each step on previous values of y. This could be a very fruitful framework for infectious disease forecast evaluation but we do not pursue it is this work.

definition of $\mathcal{X}$. For example, having more available resources usually allows for actions incurring less loss.

3. Assign F the score $s(x^F, y)$, that is, the numerical quantification via s of the consequence under outcome y of accepting x^F as a solution to the decision problem.

Thus, given definitions of $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{C}$, and a loss function s, we produce a scoring rule $S(F, y) := s(x^F, y)$ that can be applied to any F for any outcome $y \in \mathcal{Y}$. We emphasize again a key feature of this formulation: the scoring rule S is defined entirely in terms of what we take to be possible actions, states, and consequences along with a function s that assigns values to the possible consequences.

Another key feature of this procedure is that the scoring rule S it yields is proper by definition. This is because a forecaster that believes that Y has distribution F and understands the decision problem and loss function s must also believe that x^F is the optimal action to take. This, in turn, implies the belief that there is no forecast that will give a better expected score than F. Therefore, there is no incentive to reveal a forecast at odds with a forecaster's beliefs about Y.

Remark. For there to be an incentive *not* to reveal a forecast other than F we would need a *strictly* proper scoring rule, which would require more conditions on the decision problem and/or s. In the literature, a functional $T(F)$ (such as the mean or median of F) is called *elicitable* if there is a strictly proper scoring rule $S(F, y) = s(x^F, y)$ for which $T(F) = x^F$. But the focus in work on elicitability often seems to be the functional T as a starting point rather than the decision problem, which is our primary concern here.

The following sections explain in detail how the QS, WIS, CRPS (for various parameters and weightings) as well as extensions to higher dimnsional setting can be seen as outputs of our procedure.

3   Step 1: Defining decision problems and loss functions.

We begin by imagining a public health worker trying to decide at time t_0 what quantity x to procure of a protective resource such as beds, staff, or medicine in anticipation of y new infections that will be observed (and create resource demand) at time $t_1 > t_0$. The resource has a unit-cost of C, and we assign a loss of L to the consequence of a sick individual not having access to the resource. This specifies a loss function

$$s_Q(x, y) = Cx + L(y - x)_+ \tag{1}$$

where $u_+ = \max(u, 0)$. We use the subscript Q because it will be shown below that this loss function yields the quantile score for the the probability level $\alpha = 1 - C/L$. Loss functions of this general type are referred to as *piece-wise linear*, distinguishing them from smooth non-linear loss functions such as squared error. The underlying decision problem here consists of action and outcome sets $\mathcal{X}, \mathcal{Y}$ both equal to $\mathbb{R}_+$, the non-negative reals. (We ignore for now any discreteness of the resource.) The consequence set $\mathcal{C}$ is abstractly a set of over-expenditure regrets and numbers of sick individuals with unmet resource need, but the loss function s maps $\mathcal{C}$ into $\mathbb{R}_+$. We also from now on make the assumption that $C < L$ since otherwise – at least according to this loss function – there is no "cost-effectiveness" in trying to prevent even a single case of unmet need. This would imply a Bayes act x^F that is identically 0, telling us nothing about the value of F.

This decision problem may however be complicated by uncertainty at time t_0 about the unit-cost of the resource. We take such uncertainty as requiring the worker to commit in advance to procurement

levels x_i for a range of potential costs $C_i < L$ where $i \in 1, \ldots, m$. Even though only one of the x_i will end up needing to be procured, there does not seem to be any intrinsically decision theoretic principle by which our loss function should ignore counter-factual but exactly specified consequences of the decision under unrealized costs. If we wanted to commit to a *locality* principle for prediction where losses and scores only depend on predictive densities in a neighborhood of the relevant observation, this attempt would falter, and we would need to frame our methods around the log score and it's higher order relatives. But we do not see that as an appropriate principle in public health resourse allocation where probabilistic forecasts can be valuable by pushing decision makers in the "right direction" even when they assign low probability to observed outcomes. We instead adopt a "sensitivity-to-distance" principle from which piece-wise linear loss functions and the QS can be shown to naturally emerge. It is this principle that makes the loss under alternative costs seem relevant. **APG:[These are obviously incomplete thoughts but I feel like there is something here that I've been grasping for for a lng time.]**

A simple way to take account of the consequences of actions under counterfactual costs is to use the mean of all losses under certainty about the cost:

$$s_{\text{WIS}}(x, y) = \frac{1}{m} \sum_{i=1}^{m} C_i x_i + L(x_i - y)_+ \quad (C_i < L \text{ for all } i) \tag{2}$$

where we here use x to express the decision vector $(x_1, \ldots, x_m)$. Or if there are prior beliefs about the probabilities p_i of the C_i, we could use a weighted sum

$$s_{w\text{WIS}}(x, y) = \frac{1}{m} \sum_{i=1}^{m} p_i (C_i x_i + L(x_i - y)_+). \tag{3}$$

As for Q, the WIS and $wWIS$ subscripts indicate that, for particular values of the $\{C_i\}$ and L, this loss function produces the WIS or a version that weights intervals (i.e. the average of QS's for C_i and $L - C_i$) with non-standard weights (combining p_i and C_i/L).

We can also consider C a continuous random variable and accommodate this by letting $m \to \infty$ in (3) which (with some mathematical care) converts the loss into an integral

$$s_{\text{CRPS}}(x, y) = \frac{1}{L} \int_0^L (cx(c) + L(x(c) - y)_+) dc \tag{4}$$

over all relevant costs $c \in [0, L]$, or a weighted version

$$s_{w\text{CRPS}}(x, y) = \int_0^L (cx(c) + L(x(c) - y)_+) f_C(c) dc \tag{5}$$

that uses a prior density f_C on the possible costs. These loss functions will yield the traditional CRPS and the wieghted version of Gneiting and Ranjan [2011].

It is important to note that these generalizations change not only the loss function but also the structure of the decision problem by using higher dimensional action sets $\mathcal{X} = \mathbb{R}_+^m$ or the set $\mathcal{X} = \{x(c) : [0, \infty) \to \mathbb{R}_+\}$ of $\mathbb{R}_+$ valued functions on $[0, \infty)$.

We can also expand our decision problem to include multiple actions that will all, in fact (not counterfactually), have to be taken. This includes the motivating scenario for our work, the simultaneous procurement of resources across several locations (U.S. states and territories) during a pandemic. And again, summing the losses (1) over the locations (for a particular C and L) creates a loss function that

is reactive to the overall procurement decision. Let the locations be indexed by $l \in 1, \ldots, N$. The action sets for the decision problem corresponding to a single certain cost C, a discrete cost distribution over $\{C_i\}, i = 1, \ldots, m$, and a continuous cost distribution f_C with support on $[0, L]$ are now

$$\mathcal{X}_{cert} = \mathbb{R}_+^N \tag{6}$$

$$\mathcal{X}_{disc} = (\mathbb{R}_+^m)^N \tag{7}$$

$$\mathcal{X}_{cont} = \{x(c) : [0, \infty) \to \mathbb{R}_+^N\}. \tag{8}$$

And for these "vector-action" problems we can define *total* versions of our loss functions

$$s_{tQ}(x, y) = \sum_{l=1}^{N} C x_l + L(y_l - x_l)_+ \tag{9}$$

$$s_{tWIS}(x, y) = \frac{1}{m} \sum_{l=1}^{N} \sum_{i=1}^{m} C_i x_{l,i} + L(x_{l,i} - y_l)_+ \tag{10}$$

$$s_{twWIS}(x, y) = \frac{1}{m} \sum_{l=1}^{N} \sum_{i=1}^{m} p_i (C_i x_{l,i} + L(x_{l,i} - y_l)_+) \tag{11}$$

$$s_{tCRPS}(x, y) = \frac{1}{L} \sum_{l=1}^{N} \int_0^L (c x_l(c) + L(x_l(c) - y_l)_+) dc \tag{12}$$

$$s_{twCRPS}(x, y) = \sum_{l=1}^{N} \int_0^L (c x_l(c) + L(x_l(c) - y_l)_+) f_C(c) dc. \tag{13}$$

A key point for our presentation will be that the same C, set $\{C_i\}$, probabilities $\{p_i\}$, and density f_C are used for all locations simultaneously. We believe that our methods could be adapted to variations across locations (which could be important in practice), but we leave this generalization for later work.

The loss functions (and corresponding decision problems) defined so far have all been thouroughly covered, from one perspective or another, in various strands of forecast evaluation literature. We now introduce one that, while certainly familiar in operations research and related fields, has not, to our knowledge, been explicitly discussed in a forecast evaluation context.

Consider again the simultaneous resource procurement decision a public health worker must make in attempting to mitigate a pandemic across multiple locations. We now add a constraint K on the total amount of the resource available for allocation across all locations. This presents the worker with a new decision problem we call the *allocation problem* (AP or AP(K)) with a smaller action set $\mathcal{X}_K = \{x \in \mathbb{R}_+^N | \sum x_l \leq K\} \subset \mathcal{X}_{cert}\}$ from which to choose (and analogously restricted version of the sets $\mathcal{X}_{disc}$ and $\mathcal{X}_{cont}$ for the decision problems with uncertain cost). Loss functions with the basic $Cx + L(y - x)_+$ structure we have used up until now continue to generate meaningful scoring rules with respect to the AP following our procedure. But with a view toward developing a scoring rule that accommodates uncertainty about both C and K in a non-redundant (and more interpretable and computationally feasible) way, we augment the action set of the AP to include, in addition to the procurement levels x_l for all locations, a "reserve" quantity z_K defined as the available amount of resource that the worker decides not to allocate anywhere. In optimization theory quantities defined in this manner are known as *slack* variables, and are introduced in order to re-express contraint

inequalities as equalities. In our case, we get

$$\mathcal{X}_K = \{(x, z) \in \mathbb{R}_+^{N+1} \mid \sum_{l=1}^{N} x_l + z = K\}. \tag{14}$$

In these action set variables, the loss function we used for the unconstrained N location decision problem with action set $\mathcal{X}_{cert}$ now has the expression

$$s_{tQ}((x, z), y) = C(K - z) + \sum_{l=1}^{N} L(y_l - x_l)_+. \tag{15}$$

This isolates – at least formally – the explicit contribution of the procurement part, x, of the decision to the second term which we call the *allocation loss function*:

$$s_{A,K}(x, y) := L \sum_{l=1}^{N} (y_l - x_l)_+. \tag{16}$$

It will turn out, as explained in the following sections, that for a given forecast F, a scoring rule derived with respect to the AP from s_{tQ} under uncertainty about C – that is, from s_{tWIS} or s_{tCRPS} – will have the same values for all outcomes y as a scoring rule that is a weighted average or integration of scoring rules derived from allocation losses $s_{A,K}$ with respect the AP's for a range of K's.

4   Step 2: Derive the Bayes act for a forecast F

I think we should not hide the formula

$$\frac{d}{dx_{il}} E_F[s_.(x, y)] = E_F[\frac{d}{dx_{il}} s_.(x, y)] = C - L E_{F_l}[\mathbf{1}(y_l > x_{il})] = C + L(F_l(x_{il}) - 1) \tag{17}$$

$$\implies x_{il}^{F_l} = F_l^{-1}(1 - C/L) \tag{18}$$

$$\text{or for CRPS} \tag{19}$$

$$x_l^{F_l}(c) = F_l^{-1}(1 - c/L) \tag{20}$$

Fundamental point:

- When $z_K > 0$, K has no influence on the AP(K) Bayes act, i.e., $\partial_K x^{F,C,K} = 0$.

- When $z_K = 0$ and x^F are not $1 - C/L$ quantiles, $\partial_C x^{F,C,K} = 0$.

Picture: For K fixed, $x^{F,K}(c)$ will be a constant point $x^{F,K}$ on $\{\sum x_l = K\}$ for $c \in [0, c_1]$, and then wander back to 0 for $c \in (c_1, L]$. When C is fixed, $x^F(k)$ will be a point on $\{\sum x_l = k\}$ (which does not depend on C) for $k \in [0, \sum F_l^{-1}(1 - C/L)]$, and then constant at $\mathbf{F}^{-1}(1 - C/L)$ for $k \in (\sum F_l^{-1}(1 - C/L), \infty)$.

5   Step 3: Use the Bayes act to score the forecast F against outcomes.

Deal with alloscore by noticing that simultaneous quantile loss has same minimizer as alloscore with sum of quantiles as constraint level.

$$s_{t\text{WIS}}(x^{F,t\text{WIS}}, y) = \frac{1}{m} \sum_{l=1}^{N} \sum_{i=1}^{m} s_Q(x_{l,i}^{F_l,Q,C_i}, y_l)_+ \tag{21}$$

$$= \frac{1}{m} \sum_{i=1}^{m} s_{\text{AS}}(x^{F,\text{AS},K(C_i,F,L)}, y) \tag{22}$$

$$s_{\text{IAS}}(x^{F,\text{IAS}}(k), y) = \int_0^\infty s_{\text{AS}}(x^{F,\text{AS}(k)}, y) dk \tag{23}$$

$$= \int_0^\infty \sum_{l=1}^{N} s_Q(x^{F_l,Q(c(k,F))}, y) dk \tag{24}$$

$$= \sum_{l=1}^{N} \int_0^\infty s_Q(x^{F_l,Q(c(k,F))}, y) dk \tag{25}$$

But

$$k = k(c, F) = \sum_{l=1}^{N} F_l^{-1}(1 - c/L) \tag{26}$$

$$\frac{\partial k}{\partial c} = \sum_{l=1}^{N} \frac{-1}{L f_l(q_{F_l,\tau(c)})} := -\text{MAI}(c, F) \tag{27}$$

(where we assume that the F_l have non-negative support). The name MAI refers to this derivative being the marginal total allocation increase for the total quantile loss Bayes act $x^{F,Q(C)}$ that results from a reduction of the cost C of the resource by 1. Or in terms of the probability level $\tau(c) = 1 - c/L$,

$$\frac{\partial k}{\partial \tau} = \frac{\partial k}{\partial c} \frac{dc}{d\tau} = -\text{MAI}(c, F) \cdot (-L\tau) = L\text{MAI}(c, F) \tag{28}$$

Now a constraint $K = 0$ corresponds to cost-loss parity $C = L$, i.e. $\tau = 0$, for which the Bayes act is the zero vector. And $K \to \infty$ corresponds to the cost of over-allocation vanishing, i.e. $\tau \to 1$. So making a change of variables with (27) or (28) we get

$$s_{\text{IAS}}(x^{F,\text{IAS}}(k), y) = \sum_{l=1}^{N} \int_0^L s_Q(x^{F_l,Q(c)}, y)\text{MAI}(c, F) dc \tag{29}$$

$$= \sum_{l=1}^{N} \int_0^1 s_Q(x^{F_l,\tau}, y)L\text{MAI}(L(1 - \tau), F) d\tau \tag{30}$$

$$= tw\text{CRPS}(F, y) \tag{31}$$

where MAI provides the (identical) weighting in each coordinate. Crucially though, this weighting *depends* on F so that IAS's of two forecasts F and $\tilde{F}$ *cannot* be interpreted as result of applying the same wCRPS to the two forecasts.

More about SR's:

- QS: meteorologist loss v newsvendor

- WIS/CRPS: multiple scenario loss (Gruschka-Cockayne et al, Fissler and Ziegel, Jose and Winkler) - Least squares analogy

- bring in Brier?

- take stab at explaing log score following Dawid with Bayes act being inverse of $x^F(c)$ i.e., F itself.

5.1 Solving the Allocation Problem

As the solution to a constrained optimization problem, the optimal allocation x^F according to a forecast F can be described formally as the first component of a stationary point $(x^F, z_K, \lambda^\star)$ in the region $\{\mathbf{x} \geq 0, \lambda \geq 0\}$ for the Lagrangian

$$\mathcal{L}(\mathbf{x}, \lambda; K, C, L) = L \sum_{l=1}^{N} E_{F_l}[(Y_l - x_l)_+] + C(K - z_K) + \lambda \left(z_K + \sum_{l=1}^{N} x_l - K \right). \tag{32}$$

That is, optimality at x^F requires that

$$0 = \frac{\partial}{\partial x_l} \mathcal{L}(x^F, z_K, \lambda) = L E_F \left[\frac{d}{dx_l}(Y_l - x_l^F)_+ \right] + \lambda \tag{33}$$

$$= L E_F \left[-\mathbf{1}\{Y_i \geq x_i^F\} \right] + \lambda \tag{34}$$

$$= L(F_i(x_l^F) - 1) + \lambda \tag{35}$$

for any $x_l^F > 0$;

$$0 = \frac{\partial}{\partial z_K} \mathcal{L}(x^F, z_K, \lambda) = \lambda - C \tag{36}$$

if $z_K > 0$; and

$$0 = \frac{\partial}{\partial \lambda} \mathcal{L}(\mathbf{x}^F, \lambda^\star) = \sum_{i=1}^{N} x_i^F - K \tag{37}$$

$$= \sum_{\{i \mid x_i^F > 0\}} F_i^{-1}(1 - \lambda^\star) - K \quad \text{(by (35))}. \tag{38}$$

These equations give formulae $x_i^F = F_i^{-1}(1 - \lambda^\star)$, but in many situations of interest for us the F_i^{-1} are not single-valued, and even when they are, a solution $\lambda^\star$ of (38) can only be obtained analytically in special cases (but see ...). We therefore typically calculate allocation scores using approximately optimal allocations found via a binary search for $\lambda^\star$. This search begins with the interval $\lambda \in [0, 1]$ and checks at each step the sign of the inequality that results from substituting the interval midpoint λ_τ into (38) for endpoints of the sets $F_i^{-1}(1 - \lambda_\tau)$. If $\sum_i F_i^{-1}(1 - \lambda^\star) > K$, the quantiles for $1 - \lambda_\tau$ exceed our resource constraint so we set λ_τ as the lower endpoint of the $\tau + 1$ search interval. Similarly, if $\sum_i F_i^{-1}(1 - \lambda^\star) < K$, the quantiles for $1 - \lambda_\tau$ do not make full use of our resource and we set λ_τ as the upper endpoint of the $\tau + 1$ search interval. This process halves the search interval width at each step and we continue until this width falls below a given tolerance level ε_τ yielding an approximation $\lambda^{\star, \varepsilon}$. When quantiles $F_i^{-1}(1 - \lambda^{\star, \varepsilon})$ are singly defined and the forecast densities f_i not too close to zero at these quantiles, the discrepancy $\Delta(\lambda^{\star, \varepsilon}) = \sum_{i=1}^{N} F_i^{-1}(1 - \lambda^{\star, \varepsilon}) - K$ is of order ε_λ and we can take $F_i^{-1}(1 - \lambda^{\star, \varepsilon})$ as approximations to x_i^F.

Often though, $1 - \lambda^{\star, \varepsilon}$ will be the probability level of a "plateau" for one or more F_i and the λ search algorithm will fail to drive $\Delta(\lambda^{\star, \varepsilon})$ close enough to zero for the $F_i^{-1}(1 - \lambda^{\star, \varepsilon})$ to be acceptable approximations of the optimal allocations. In this event we follow-up with a "post-processing" step that

identifies lower and upper boundaries $x_{i,L}$ and $x_{i,U}$ of any such low-density plateau containing points in $F_i^{-1}(1 - \lambda^{\star,\varepsilon})$. If $F_i^{-1}(1 - \lambda^{\star,\varepsilon})$ is single-valued with the density $f_i(F_i^{-1}(1 - \lambda^{\star,\varepsilon}))$ not close to zero, we simply take $x_{i,L} = F_i^{-1}(1 - \lambda^{\star,\varepsilon}) = x_{i,U}$. Defining $\mathbf{x}_L = (x_{i,L})$ and $\mathbf{x}_U = (x_{i,U})$ we then interpolate an approximately optimal allocation $\mathbf{x} = (1 - t^\star)\mathbf{x}_L + t^\star\mathbf{x}_U$ where $t^\star$ solves $\sum_{i=1}^{N}(1 - t^\star)x_{i,L} + t^\star x_{i,U} = K$.

We note though that by using this linear interpolation (rather than some other point in the plateau satisfying the constraint) as the allocation associated with F in the scoring process we are making a choice that is not guided by any expected utility maximization principle. As such, this choice may imply strong and potentially questionable assumptions about fairness and equity principles for the decision maker when forecasts tend to be "clumped," leaving many regions with very low forecasted probability.

6   Discussion

7   References

References

Tilmann Gneiting and Roopesh Ranjan. Comparing density forecasts using threshold- and quantile-weighted scoring rules. *Journal of Business & Economic Statistics*, 29(3):411–422, 2011. doi: 10. 1198/jbes.2010.08110. URL https://doi.org/10.1198/jbes.2010.08110.