

SCORING LOGIC: PATIENT CHOICE OF HOSPITALS

BRYCE BURGESS

1. EFFECTIVE SCORING

An effective scoring method fills a number of constraints (current methods do not necessarily quite fit). On the domain, the output of the scoring method should be in $[0,1]$ (may change to $[-1,1]$). The minimum input should map to the minimum output, and the same goes for the maximum input and output.

The slope must be monotonic: all mappings to new values should be unique. Whether the slope is positive or negative will be defined on a per-variable basis

If a patient's choice and the measure of center on the hospital are the same (or close), the output score should reflect that it was neutral.

If mostly low scores are available from the hospitals, a uniform patient distribution should be reweighted to produce more high-scoring patients for that variable. If mostly high scores are available from the hospitals, a uniform patient distribution should be reweighted to produce more low-scoring patients for that variable.

v	any hospital variable, the value of a hospital variable taken from the hospital that the patient chose (e.g. num
P	set of all patients and their associated data such as age, sex, ruralness, hospital choice,
p	individual patient in P
$p(v)$	value of variable v for patient j
$p(hospital)$	chosen hospital for patient j
p_h	chosen hospital for patient j
$P^{[0,1]}(v)$	patient variable values after compressing them to the range $[0, 1]$
H	set of all hospitals and their associated data
$H(v)$	value of hospital variable v
h	individual hospital i and associated data
H_p	list of hospitals to choose from for patient j
$\overline{H_p(v)}$	mean of hospitals to choose from for variable v and patient j
$H_p^{[0,1]}(v)$	hospital mean of variable after compressing them to the range $[0, 1]$
$\alpha(t)$	multiplier dependent on a travel time t
$T(p, h)$	travel time from a plz code to a given hospital
$max_{eff}(S)$	95 th percentile value of set S

2. SELECTING HOSPITALS

2.1. **Option 1: include all.** include all hospitals

$$H_p = H \text{ for every patient } p$$

2.2. **Option 2: constant multiplier to chosen.** H_p includes h if time from plz to h is less than α_0 times the distance to the chosen hospital:

$$\alpha(t) = \alpha_0 t$$

where α_0 is a constant.

$$H_p = \{h \mid h \in H \text{ s.t. } T(p, h) < \alpha_0 T(p, p_h)\} \text{ options for patient } p$$

2.3. **Option 3: proportional multiplier to chosen.** H_p includes h if time from plz to hospital is less than $\alpha(t)$ times the distance to the chosen hospital

$$\alpha(t) = \log(t)t$$

$$H_p = \{h \mid h \in H \text{ s.t. } T(p, h) < \log(T(p, p_h))T(p, p_h)\} \text{ options for patient } p$$

2.4. **Option 4: chosen raised to a constant.** H_p includes h if time from plz to h is less than the distance to the chosen hospital raised to the power α_0 :

$$\alpha(t) = t^{\alpha_0}$$

where α_0 is a constant.

$$H_p = \{h \mid h \in H \text{ s.t. } T(p, h) < T(p, p_h)^{\alpha_0}\} \text{ options for patient } p$$

3. LINEAR SCORING

Linearly transform each patient value to fit within the range $[0,1]$, based on the entire set of patient values. Additionally, anything above \max_{eff} is set to 1, as a way to mitigate the effects of outliers. In this way the set $P^{[0,1]}(v)$ is equal to:

$$P^{[0,1]}(v) = \left\{ \min\left(\frac{p(v) - \min(P(v))}{\max_{eff}(P(v))}, 1\right) \right\}$$

Linearly transform the hospital mean in the same way as the patient set, so they can still be meaningfully compared.

$$\overline{H_p^{[0,1]}(v)} = \frac{\overline{H_p(v)} - \min(P(v))}{\max_{eff}(P(v))}$$

If higher numbers are interpreted as more important, apply the exponent below, which is scaled by $\overline{H_p^{[0,1]}(v)}$. The logarithm base in the expression $\frac{\log(0.5)}{\log(\overline{H_p^{[0,1]}(v)})}$ doesn't matter, as it equivalent to $\log_{\overline{H_p^{[0,1]}(v)}}(0.5)$ by the logarithm change of base formula:

$$P^{score}(v) = \left\{ p^{[0,1]}(v)^{\frac{\log(0.5)}{\log(\overline{H_p^{[0,1]}(v)})}} \right\}$$

If lower numbers are interpreted as more important, and the change is the same as before, except that all $p^{[0,1]}(v)$ and $\overline{H_p^{[0,1]}(v)}$ are inverted within the range $[0,1]$ ²:

$$P^{score}(v) = \left\{ \left(1 - p^{[0,1]}(v) \right)^{\frac{\log(0.5)}{\log(1 - \overline{H_p^{[0,1]}(v)})}} \right\}$$

	$\overline{H_p^{[0,1]}(v)} = 0.25$	$\overline{H_p^{[0,1]}(v)} = 0.5$	$\overline{H_p^{[0,1]}(v)} = 0.75$
$p^{[0,1]}(v) = 0.25$	0.5	0.25	0.035
$p^{[0,1]}(v) = 0.5$	0.707	0.5	0.188
$p^{[0,1]}(v) = 0.75$	0.866	0.75	0.5

TABLE 1. table of patient choice values and hospital means. Note that when the hospital mean is equal to 0.5, each patient value is unchanged, and any time that the patient value is equal to the hospital mean, the output score is 0.5, showing that the variable had no priority for that individual patient

4. EXPONENTIAL SCORING

(1) if $\overline{H_p(v)} = 0$ set it to 0.01

(2) if lower numbers are interpreted as better:

(3) $P^{score}(v) = \left\{ 2^{\left(-\frac{p(v)}{\overline{H_p(v)}} \right)} \right\}$

(4) otherwise:

(5) $P^{score}(v) = \left\{ 1 - 2^{\left(-\frac{p(v)}{\overline{H_p(v)}} \right)} \right\}$

	$\frac{p(v)}{\overline{H_p(v)}} = 0.5$	$\frac{p(v)}{\overline{H_p(v)}} = 1$	$\frac{p(v)}{\overline{H_p(v)}} = 2$
$P^{score}(v)$	0.25	0.5	0.71

TABLE 2. table of output scores for variables that are positively scaling (bigger is better), given some possible ratios of inputs. These are not compressed as in linear scoring, so both $p(v)$ and $\overline{H_p(v)}$ cover a very large possible domain.

¹In the case that $\overline{H_p^{[0,1]}(v)}$ is equal to zero, instead it is set equal to a near-zero positive real number whereas if $\log(\overline{H_p^{[0,1]}(v)})$ is equal to zero, it is set equal to a near-zero negative real number.

²In the case that $1 - \overline{H_p^{[0,1]}(v)}$ is equal to zero, instead it is set equal to a near-zero positive real number whereas if $\log(1 - \overline{H_p^{[0,1]}(v)})$ is equal to zero, it is set equal to a near-zero negative real number.