Institute for Health
Metrics and Evaluation

Joint Statistical Meeting, August 2024

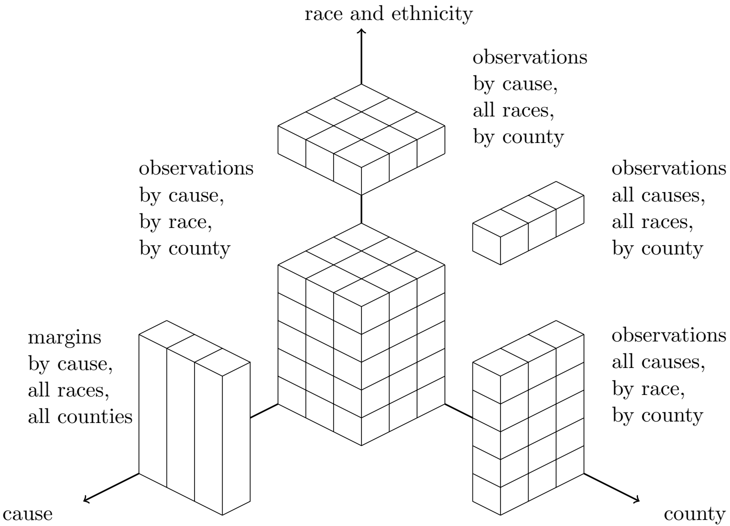
Domain-specific raking, with application to mortality modeling

Ariane Ducellier, Alexander Hsu, Aleksandr Aravkin, Peng Zheng

Outline

- Motivating example
- Raking problem
- Uncertainty propagation
- Application

Motivating example



Our problem

The constraints are the marginal totals s of the observations y_i:

- In 1D, $\sum_i \beta_i = s$
- In 2D, $\sum_i \beta_{i,j} = s_{1,j}$ and $\sum_j \beta_{i,j} = s_{2,i}$

Our objective is to find the raked values β_i respecting the constraints. The general problem can be written as:

$$\min_{\beta} f(\beta; y) \quad \text{s.t.} \quad A\beta = s$$

For instance, if we choose the entropic distance for f, we have:

$$\min_{\beta} \sum_i \beta_i \log \frac{\beta_i}{y_i} - \beta_i + y_i \quad \text{s.t.} \quad A\beta = s$$

Primal-dual formulation

Primal problem:

$$\min_{\beta} \max_{\lambda} \mathcal{L}(\beta, \lambda) := f(\beta; y) + \lambda^T (A\beta - s)$$

Dual problem:

$$\min_{\lambda} \lambda^T s + f^* \left(-A^T \lambda\right) \quad \Rightarrow \quad \min_{\lambda} \lambda^T s + y^T \exp\left(-A^T \lambda\right)$$

where f^* is the convex conjugate:

$$f^*(z) = \sup_{\beta} z^T \beta - f(\beta; y)$$

Solving the dual problem

We have a relationship between the gradient maps: $\nabla_z f^* (\cdot) = \nabla_\beta f (\cdot)^{-1}$.

The solution of the dual problem is found by solving for λ the non-linear system:

$$s - A\nabla_z f^* \left(-A^T\lambda^*; y\right) = 0 \quad \Rightarrow \quad s - A\left[y \odot \exp\left(-A^T\lambda\right)\right] = 0$$

We solve this system using Newton's method.

The raked values are then equal to:

$$\beta^* = \nabla_z f^* \left(-A^T\lambda^*\right) \quad \Rightarrow \quad \beta^* = y \odot \exp\left(-A^T\lambda^*\right)$$

Distances

$$\mathcal{L}_2 \text{ norm: } f(\beta; y) = \sum_{i=1}^{p} \frac{1}{2}(\beta_i - y_i)^2$$

$$\chi^2 \text{ distance: } f(\beta; y) = \sum_{i=1}^{p} \frac{1}{2y_i}(\beta_i - y_i)^2$$

$$\text{Entropic distance: } f(\beta; y) = \sum_{i=1}^{p} \beta_i \log\left(\frac{\beta_i}{y_i}\right) - \beta_i + y_i$$

$$\text{Logit distance: } f(\beta; y) = \sum_{i=1}^{p} (\beta_i - l_i) \log \frac{\beta_i - l_i}{y_i - l_i} + (h_i - \beta_i) \log \frac{h_i - \beta_i}{h_i - y_i}$$

Uncertainties

Given:

- Σ_y the $p \times p$ covariance matrix of the observations vector y,
- Σ_s, the $k \times k$ covariance matrix of the aggregate observations vector s and
- Σ_{ys} the $p * k$ covariance matrix of y and s,

find:

- Σ_{β^*} the $p \times p$ covariance matrix of the estimated raked values β^*.

Primal problem

The primal problem:

$$F(\beta, \lambda; y, s) = \begin{bmatrix} \nabla_\beta f(\beta; y) + A^T \lambda \\ A\beta - s \end{bmatrix} = 0$$

has solution:

$$\beta^* = \phi(y; s) \text{ with } \phi : \mathbb{R}^{p+k} \to \mathbb{R}^p$$

Central limit theorem

n vectors of observations $\begin{pmatrix} y^{(i)} \\ s^{(i)} \end{pmatrix}$ taking values in $\mathbb{R}^{p+k}$ with sample mean $\begin{pmatrix} \bar{y}_n \\ \bar{s}_n \end{pmatrix}$.

$$\sqrt{n}\left(\begin{pmatrix} \bar{y}_n \\ \bar{s}_n \end{pmatrix} - \theta\right) \to \mathcal{N}\left(0, \Sigma\right)$$

where:

- $\theta = \mathbb{E}\begin{pmatrix} y^{(i)} \\ s^{(i)} \end{pmatrix}$,
- Σ is the covariance matrix $\begin{pmatrix} \Sigma_y & \Sigma_{ys} \\ \Sigma_{ys} & \Sigma_s \end{pmatrix}$ and
- $\mathcal{N}\left(0, \Sigma\right) = T$ denotes the MVN distribution with mean 0 and covariance matrix Σ.

Delta method

$$\sqrt{n}\left(\phi\left(\begin{pmatrix}\bar{y}_n \\ \bar{s}_n\end{pmatrix}\right) - \phi(\theta)\right) \to \mathcal{N}\left(0, \phi'_\theta(T)\,\Sigma\phi'^T_\theta(T)\right)$$

where:

$$\phi'_\theta(T) = \begin{pmatrix} \frac{\partial\phi_1}{\partial y_1}(\theta) & \dots & \frac{\partial\phi_1}{\partial y_p}(\theta) & \frac{\partial\phi_1}{\partial s_1}(\theta) & \dots & \frac{\partial\phi_1}{\partial s_k}(\theta) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial\phi_p}{\partial y_1}(\theta) & \dots & \frac{\partial\phi_p}{\partial y_p}(\theta) & \frac{\partial\phi_p}{\partial s_1}(\theta) & \dots & \frac{\partial\phi_p}{\partial s_k}(\theta) \end{pmatrix}$$

Linearity of the solution for the χ^2 distance

$$F(\beta, \lambda; y, s) = 0 \quad \Rightarrow \quad \begin{pmatrix} I & [\text{diag}\,(y)]\, A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \beta \\ \lambda \end{pmatrix} = \begin{pmatrix} y \\ s \end{pmatrix}$$

If:

- β_1^* is a solution of $F(\beta, \lambda; y_1, s_1) = 0$ and
- β_2^* is a solution of $F(\beta, \lambda; y_2, s_2) = 0$

then:

- $\beta_1^* + \beta_2^*$ is a solution of $F(\beta, \lambda; y_1 + y_2, s_1 + s_2) = 0$ and
- $\alpha\beta_1^*$ is a solution of $F(\beta, \lambda; \alpha y, \alpha s) = 0$

Thus:

$$\phi\left(\begin{pmatrix} \bar{y}_n \\ \bar{s}_s \end{pmatrix}\right) = \frac{1}{n} \sum_{i=0}^{n} \phi\left(\begin{pmatrix} y^{(i)} \\ s^{(i)} \end{pmatrix}\right)$$

Central limit theorem

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=0}^{n}\phi\left(\begin{pmatrix} y^{(i)} \\ s^{(i)} \end{pmatrix}\right) - \phi\left(\theta\right)\right) \to \mathcal{N}\left(0, \Sigma_{\beta}\right)$$

Using:

$$\sqrt{n}\left(\phi\left(\begin{pmatrix} \bar{y}_n \\ \bar{s}_n \end{pmatrix}\right) - \phi\left(\theta\right)\right) \to \mathcal{N}\left(0, \phi_{\theta}'\left(T\right)\Sigma\phi_{\theta}'^{T}\left(T\right)\right)$$

we get:

$$\Sigma_{\beta} = \phi_{\theta}'\left(T\right)\Sigma\phi_{\theta}'^{T}\left(T\right)$$

$\Rightarrow$ We need to compute the partial derivatives $\frac{\partial \phi_j}{\partial y_i}$ and $\frac{\partial \phi_j}{\partial s_i}$.

Implicit Function Theorem

$F: S \in \mathbb{R}^{2p+2k} \to \mathbb{R}^{p+k}$ is a function of class C^1 and:

$$F(\beta^*, \lambda^*; y_0, s_0) = 0 \text{ and } \det D_{\beta,\lambda} F(\beta^*, \lambda^*; y_0, s_0) \neq 0$$

where:

$$D_{\beta,\lambda} F = \begin{pmatrix} \frac{\partial F_1}{\partial \beta_1} & \cdots & \frac{\partial F_1}{\partial \beta_p} & \frac{\partial F_1}{\partial \lambda_1} & \cdots & \frac{\partial F_1}{\partial \lambda_k} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial F_{p+k}}{\partial \beta_1} & \cdots & \frac{\partial F_{p+k}}{\partial \beta_p} & \frac{\partial F_{p+k}}{\partial \lambda_1} & \cdots & \frac{\partial F_{p+k}}{\partial \lambda_k} \end{pmatrix}$$

then $F(\beta, \lambda; y, s) = 0$ defines a function $(\beta, \lambda) = \phi(y, s)$ for $(y, s) \in \mathbb{R}^{p+k}$ near (y_0, s_0), with $(\beta, \lambda) = \phi(y, s)$ close to (β^*, λ^*).

Implicit Function Theorem

When differentiating $F(y, s; \phi(y, s)) = 0$ at the solution (β^*, λ^*), we get:

$$\left[D_{\beta,\lambda} F(y, s; \beta^*, \lambda^*)\right] \left[D_{y,s} \phi(y, s)\right] + \left[D_{y,s} F(y, s; \beta^*, \lambda^*)\right] = 0$$

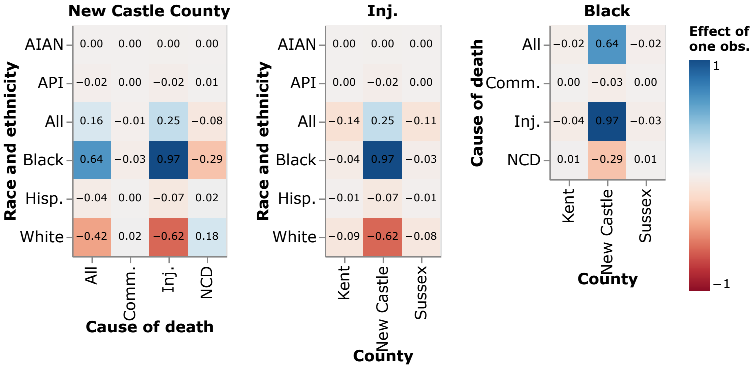
We have:

$$D_{\beta,\lambda} F = \begin{pmatrix} \nabla^2_\beta f(\beta; y) & A^T \\ A & 0_{k\times k} \end{pmatrix} \text{ and } D_{y,s} F = \begin{pmatrix} \nabla^2_{\beta y} f(\beta; y) & 0_{p\times k} \\ 0_{k\times p} & -I_{k\times k} \end{pmatrix}$$

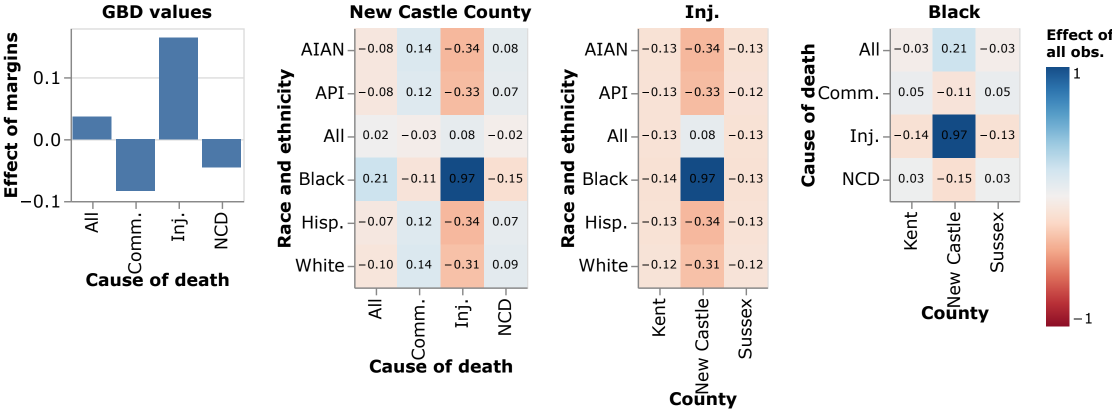
thus we can compute:

$$D_{y,s}\phi = \begin{pmatrix} \frac{\partial\beta}{\partial y} & \frac{\partial\beta}{\partial s} \\ \frac{\partial\lambda}{\partial y} & \frac{\partial\lambda}{\partial s} \end{pmatrix}$$

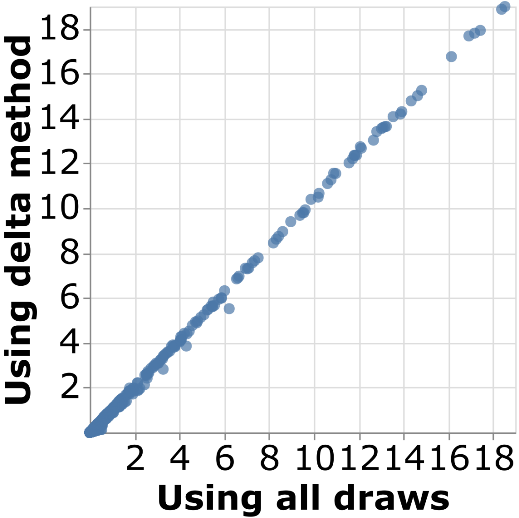
How changes in one observation will affect the raked values?



How changes in observations and margins will affect one raked value?



What are the final uncertainties on the raked values?



Questions?

https://arxiv.org/abs/2407.20520