

$$(1) \mathcal{F}_R = \left\{ m: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0} \mid \begin{array}{l} m(u) = c g_R\left(\frac{u-a}{b}\right) \text{ with } a, b, c > 0 \\ \text{and } g_R \text{ a density with mean } 0, \text{ sd } 1 \end{array} \right\}$$

$$m_i(u) = c_i g_R\left(\frac{u-a_i}{b_i}\right) \quad \forall i \in \mathcal{I}_R.$$

$$\begin{aligned} \int_{-\infty}^{\infty} m_i(u) du &= c_i \int g_R\left(\frac{u-a_i}{b_i}\right) du \\ &= c_i \int b_i g_R(t) dt = c_i b_i \end{aligned}$$

$$\begin{aligned} \int_{-\infty}^{\infty} u m_i(u) du &= c_i \int u g_R\left(\frac{u-a_i}{b_i}\right) du \\ &= c_i b_i \int (a_i + b_i t) g_R(t) dt \\ &= c_i b_i a_i \end{aligned}$$

$$\begin{aligned} \int_{-\infty}^{\infty} u^2 m_i(u) du &= c_i \int u^2 g_R\left(\frac{u-a_i}{b_i}\right) du \\ &= c_i b_i \int (a_i + b_i t)^2 g_R(t) dt \\ &= c_i b_i a_i^2 + c_i b_i b_i^2 \end{aligned}$$

Define

$$m_i^*(u) = \frac{m_i(u)}{\int_{-\infty}^{\infty} m_i(v) dv}.$$

Then

$$a_i = \int_{-\infty}^{\infty} u m_i^*(u) du$$

$$b_i = \left\{ \underbrace{\int_{-\infty}^{\infty} u^2 m_i^*(u) du - \left(\int_{-\infty}^{\infty} u m_i^*(u) du \right)^2}_{=: (+)} \right\}^{1/2}$$

$$c_i = \int_{-\infty}^{\infty} m_i(u) du / (+).$$

(2) Define

$$\hat{m}_i(u) = \frac{\sum_{k=1}^T K\left(\frac{u - t/\sqrt{T}}{h}\right) y_{it}}{\sum_{k=1}^T K\left(\frac{u - t/\sqrt{T}}{h}\right)}$$

(K rectangular, $h = 3.5/\sqrt{T}$ (7 days), $h = 7/\sqrt{T}$ (14 days), ...) and

$$\hat{m}_i^*(u) = \frac{\hat{m}_i(u)}{\int_{-\infty}^{\infty} \hat{m}_i(v) dv}.$$

Note that

$$\int_{-\infty}^{\infty} v^l \hat{m}_i(v) dv \approx \sum_{s=1}^T \frac{1}{\sqrt{T}} \left(\frac{s}{\sqrt{T}}\right)^l \hat{m}_i\left(\frac{s}{\sqrt{T}}\right). (*)$$

Then

$$\hat{a}_i = \int_{-\infty}^{\infty} u \hat{m}_i^*(u) du$$

$$\hat{b}_i = \left\{ \int_{-\infty}^{\infty} u^2 \hat{m}_i^*(u) du - \left(\int_{-\infty}^{\infty} u \hat{m}_i^*(u) du \right)^2 \right\}^{1/2}$$

$$\hat{c}_i = \int_{-\infty}^{\infty} \hat{m}_i^*(u) du / \hat{b}_i$$

where we use the approximation (*) to compute the above integrals numerically.

(3) By definition,

$$m_i(a_i + b_i v) / c_i = g_X(v) \quad \forall i \in \mathcal{G}_R.$$

Hence, compute

$$\hat{p}_i(v) = \hat{m}_i(\hat{a}_i + \hat{b}_i v) / \hat{c}_i \quad \forall i = 1, \dots, m$$

as well as

$$\hat{p}_i^*(v) = \frac{\hat{p}_i(v)}{\int_{-\infty}^{\infty} \hat{p}_i(w) dw} \quad \forall i = 1, \dots, m$$

(which is a proper density).

Now calculate the (squared) Hellinger distances

$$\Delta_{ij} = \int \left(\sqrt{\frac{1}{n_i} p_i(w)} - \sqrt{\frac{1}{n_j} p_j(w)} \right)^2 dw$$

and define the dissimilarity measure

$$D(S, S') = \max_{i \in S, j \in S'} \Delta_{ij}$$

for $S, S' \subseteq \{1, \dots, n\}$. Run a HAC algorithm with this dissimilarity measure.