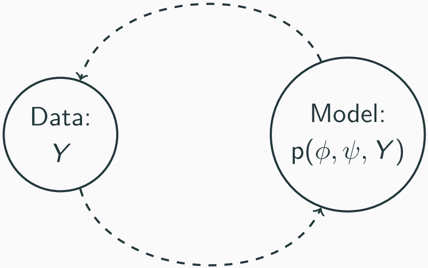
Markov melding, submodel conflict, and self-density ratios

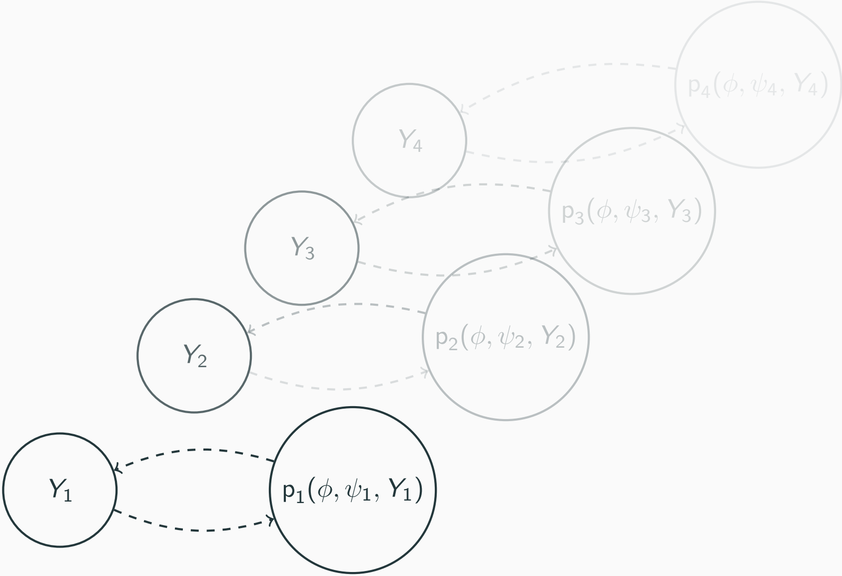
Andrew Manderson

2019-10-03

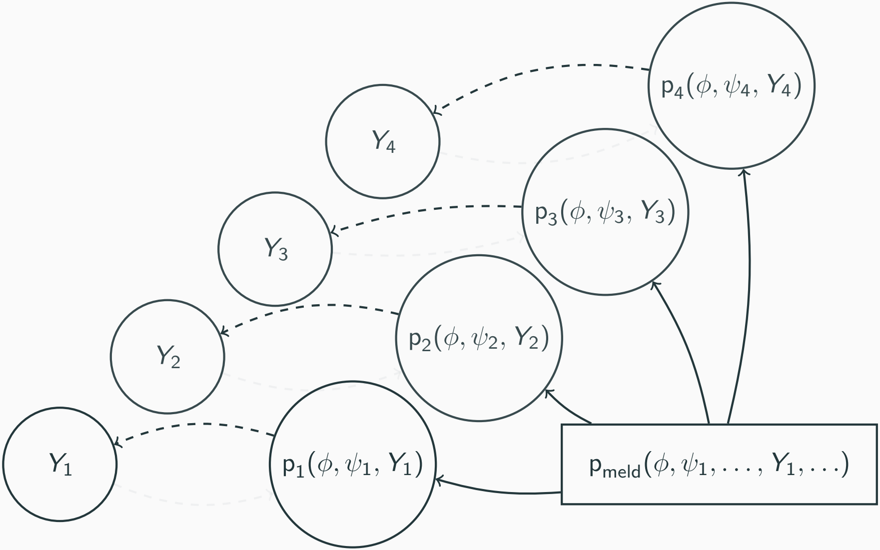
Statistical modelling



Modern statistical modelling



Melding based statistical modelling?



Markov melding

- Markov melding (Goudie et al. 2019) lets us join models with common component: ϕ
- Ideally we would specify the generative model (For $M = 2$ submodels):

$$\mathrm{p}_{\mathrm{meld}}(\phi, \psi_1, \psi_2, Y_1, Y_2) = \mathrm{p}_{\mathrm{pool}}(\phi)\ \mathrm{p}_1(\psi_1, Y_1 \mid \phi)\ \mathrm{p}_2(\psi_2, Y_2 \mid \phi)$$

- Models developed in isolation, re-express as:

$$\mathrm{p}_{\mathrm{meld}}(\phi, \psi_1, \psi_2, Y_1, Y_2) = \mathrm{p}_{\mathrm{pool}}(\phi)\ \frac{\mathrm{p}_1(\phi, \psi_1, Y_1)}{\mathrm{p}_1(\phi)}\ \frac{\mathrm{p}_2(\phi, \psi_2, Y_2)}{\mathrm{p}_2(\phi)}$$

- Analytic form of $\mathrm{p}_m(\phi)$ often unknown

Multi-stage sampler

- Need to sample the melded posterior:

$$\begin{aligned}\mathsf{p}_{\text{meld}}(\phi, \psi_1, \psi_2 \mid Y_1, Y_2) &\propto \mathsf{p}_{\text{meld}}(\phi, \psi_1, \psi_2, Y_1, Y_2) \\ &= \mathsf{p}_{\text{pool}}(\phi) \, \frac{\mathsf{p}_1(\phi, \psi_1, Y_1)}{\mathsf{p}_1(\phi)} \, \frac{\mathsf{p}_2(\phi, \psi_2, Y_2)}{\mathsf{p}_2(\phi)}\end{aligned}$$

- Submodels can be complex $\rightarrow$ Sample in stages

Stage one acceptance probability

- Say we we choose to target:

$$\mathsf{p}_{\mathsf{meld}}(\phi, \psi_1, \psi_2 \mid Y_1, Y_2) \;\propto\; \mathsf{p}_{\mathsf{pool}}(\phi)\, \frac{\mathsf{p}_1(\phi, \psi_1, Y_1)}{\mathsf{p}_1(\phi)}\, \frac{\mathsf{p}_2(\phi, \psi_2, Y_2)}{\mathsf{p}_2(\phi)}$$

- Given our stage one target, the stage one sampler has the following acceptance probability:

$$\alpha((\phi^*, \psi_1^*), (\phi, \psi_1)) = \frac{\mathsf{p}_1(\phi^*, \psi_1^*, Y_1)\mathsf{p}_1(\phi)\mathcal{Q}(\phi, \psi_1 \mid \phi^*, \psi_1^*)}{\mathsf{p}_1(\phi, \psi_1, Y_1)\mathsf{p}_1(\phi^*)\mathcal{Q}(\phi^*, \psi_1^* \mid \phi, \psi_1)}$$

- (asymptotically) Produces samples from the stage one target

Stage two acceptance probability

- Stage two target, the melded posterior:

$$\mathrm{p}_{\text{meld}}(\phi, \psi_1, \psi_2 \mid Y_1, Y_2) \propto \mathrm{p}_{\text{pool}}(\phi) \frac{\mathrm{p}_1(\phi, \psi_1, Y_1)}{\mathrm{p}_1(\phi)} \frac{\mathrm{p}_2(\phi, \psi_2, Y_2)}{\mathrm{p}_2(\phi)}$$

- Use stage one samples of ϕ as the proposal distribution in a Gibbs update for $\phi \mid \psi$:

$$\alpha(\phi^*, \phi) = \frac{\mathrm{p}_{\text{pool}}(\phi^*)\mathrm{p}_1(\phi^*, \psi_1, Y_1)\mathrm{p}_2(\phi^*, \psi_2, Y_2)\mathrm{p}_1(\phi)\mathrm{p}_2(\phi)}{\mathrm{p}_{\text{pool}}(\phi)\mathrm{p}_1(\phi, \psi_1, Y_1)\mathrm{p}_2(\phi, \psi_2, Y_2)\mathrm{p}_1(\phi^*)\mathrm{p}_2(\phi^*)} \frac{\mathrm{p}_1(\phi, \psi_1, Y_1)\mathrm{p}_1(\phi^*)}{\mathrm{p}_1(\phi^*, \psi_1, Y_1)\mathrm{p}_1(\phi)}$$

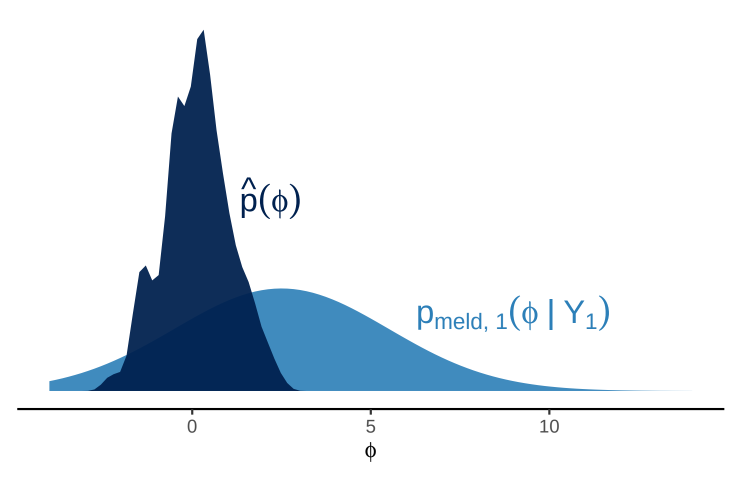
- Stage one terms cancel:

$$\alpha(\phi^*, \phi) = \frac{\mathrm{p}_{\text{pool}}(\phi^*)\mathrm{p}_2(\phi^*, \psi_2, Y_2)\mathrm{p}_2(\phi)}{\mathrm{p}_{\text{pool}}(\phi)\mathrm{p}_2(\phi, \psi_2, Y_2)\mathrm{p}_2(\phi^*)}$$

- $\mathrm{p}_2(\phi)$ not known analytically $\rightarrow$ substitute KDE: $\hat{\mathrm{p}}_2(\phi)$

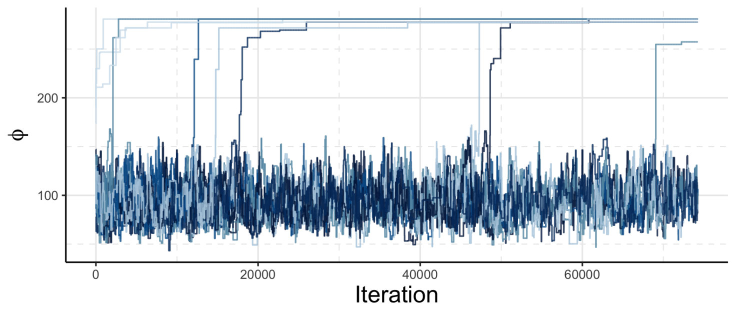
Conflict

$$\alpha(\phi^*, \phi) = \frac{\mathrm{p_{pool}}(\phi^*)\mathrm{p}_2(\phi^*, \psi_2, Y_2)}{\mathrm{p_{pool}}(\phi)\mathrm{p}_2(\phi, \psi_2, Y_2)} \cdot \frac{\hat{\mathrm{p}}_2(\phi)}{\hat{\mathrm{p}}_2(\phi^*)}$$



Conflict in action

$$\alpha(\phi^*, \phi) = \frac{\mathsf{p}_{\mathsf{pool}}(\phi^*)\mathsf{p}_2(\phi^*, \psi_2, Y_2)}{\mathsf{p}_{\mathsf{pool}}(\phi)\mathsf{p}_2(\phi, \psi_2, Y_2)} \cdot \frac{\hat{\mathsf{p}}_2(\phi)}{\hat{\mathsf{p}}_2(\phi^*)}$$



- H1N1 example from Goudie et al. (2019)

Self density ratios

- Only interact with unknown marginal via self-density ratio (Hiraoka, Hamada, and Hori 2018):

$$\mathrm{r}(\phi_{\mathrm{nu}}, \phi_{\mathrm{de}}) = \frac{\mathrm{p}(\phi_{\mathrm{nu}})}{\mathrm{p}(\phi_{\mathrm{de}})}$$

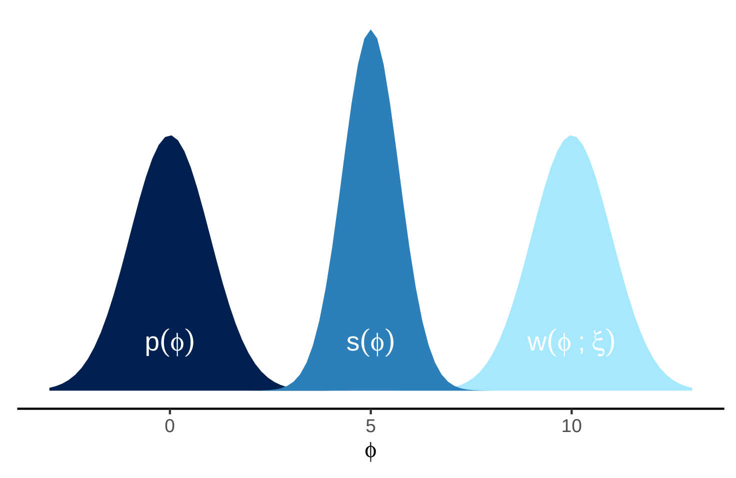
- *Weighted-sample self-density ratio estimation* (WSRE) Intuition:
  - Sampling $\phi \sim \mathrm{p}(\phi, \psi, Y)$ admits $\phi \sim \mathrm{p}(\phi)$
  - Instead, sample: $\phi \sim \mathrm{p}(\phi, \psi, Y)\mathrm{w}(\phi; \eta) \ \rightarrow \ \phi \sim \mathrm{p}(\phi)\mathrm{w}(\phi; \eta) = \mathrm{s}(\phi)$
  - Use the weighted sample density estimator of Jones (1991):

$$\hat{\mathrm{p}}(\phi) = \frac{1}{ZNh} \sum_{i=1}^{N} \mathrm{w}(\phi; \eta)^{-1} \mathrm{K}_h(\phi - \phi_i)$$

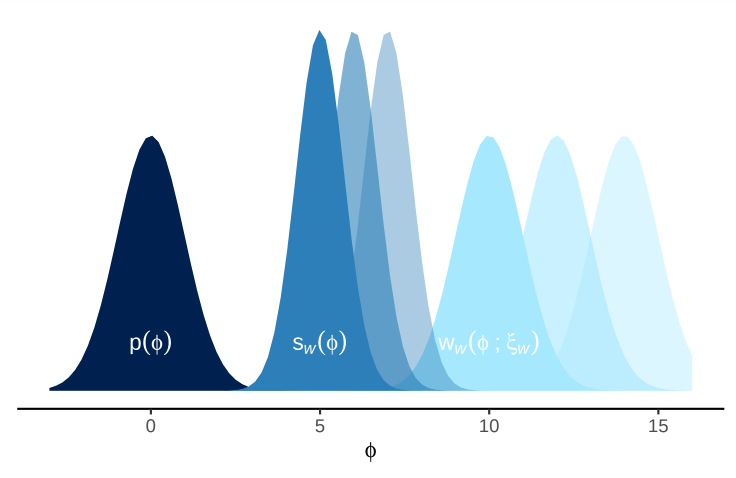
  - Ratio estimator is then:

$$\hat{\mathrm{r}}(\phi_{\mathrm{nu}}, \phi_{\mathrm{de}}) = \frac{\hat{\mathrm{p}}(\phi_{\mathrm{nu}})}{\hat{\mathrm{p}}(\phi_{\mathrm{de}})}$$

WSRE intuition - 1

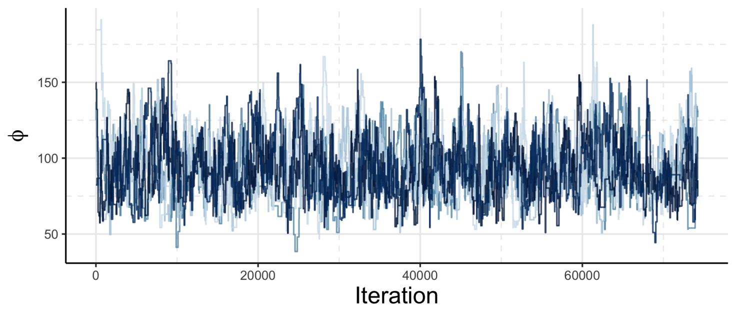


WSRE intuition - W



Fixed!

$$\alpha(\phi^*, \phi) = \frac{\mathrm{p_{pool}}(\phi^*)\mathrm{p_2}(\phi^*, \psi_2, Y_2)}{\mathrm{p_{pool}}(\phi)\mathrm{p_2}(\phi, \psi_2, Y_2)} \cdot \hat{\mathrm{r}}(\phi, \phi^*)$$



References

Goudie, Robert J. B., Anne M. Presanis, David Lunn, Daniela De Angelis, and Lorenz Wernisch. 2019. "Joining and Splitting Models with Markov Melding." *Bayesian Anal.* 14 (1). International Society for Bayesian Analysis: 81–109. doi:10.1214/18-BA1104.

Hiraoka, Kazuyuki, Toshihiko Hamada, and Gen Hori. 2018. "Necessary and Sufficient Conditions of Proper Estimators based on Self Density Ratio for Unnormalized Statistical Models." *Neural Networks* 98: 263–70. doi:10.1016/j.neunet.2017.11.018.

Jones, M. C. 1991. "Kernel Density Estimation for Length Biased Data." *Biometrika* 78 (3). Oxford University Press, Biometrika Trust: 511–19. http://www.jstor.org/stable/2337020.