Information about ϕ: Mysterious, inconvenient, or wrong results from real datasets

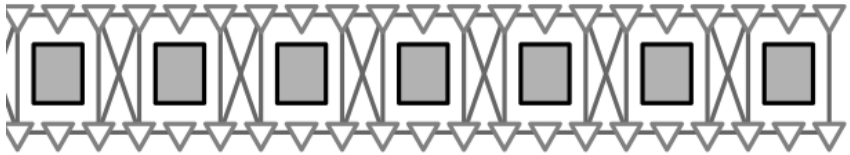
I'll show five examples of mysterious, inconvenient, or plainly wrong results produced using mixed linear models to analyze real datasets.

All involve the data's information about ϕ, the unknowns in $\mathbf{R}$ or $\mathbf{G}$.

This leads to examination of the restricted likelihood and its close relative, the marginal posterior of ϕ.

1. Periodontal data and the ICAR model

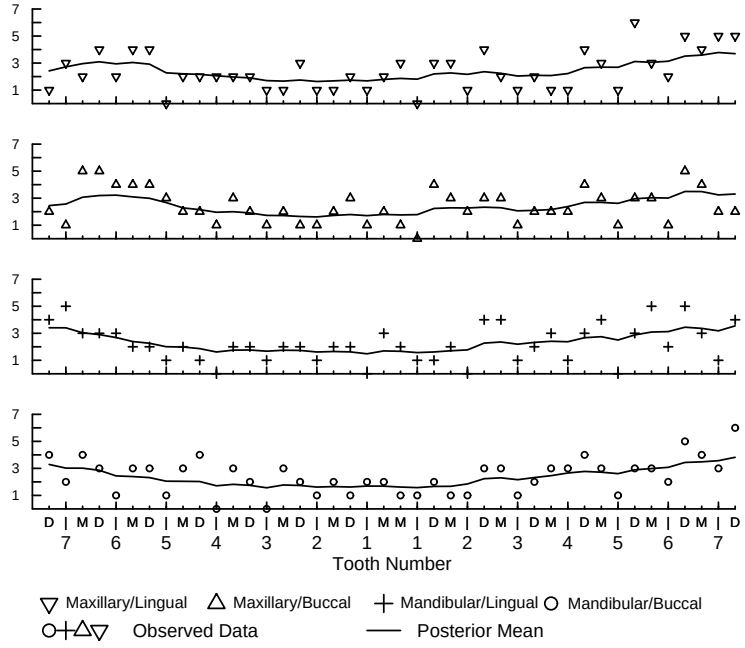
We analyzed one patient's CAL measurements using an ICAR with these neighbor pairs.



Priors: Flat on the two island means; σ_e^2 and $\sigma_s^2 \sim$ IG(0.01,0.01).

Posterior medians: σ_e^2 1.25, σ_s^2 0.25, σ_e^2/σ_s^2 4.0.

Here are the site-specific posterior means – highly smoothed.



Observation: Direct sites tend to have lower CAL than interproximal sites.

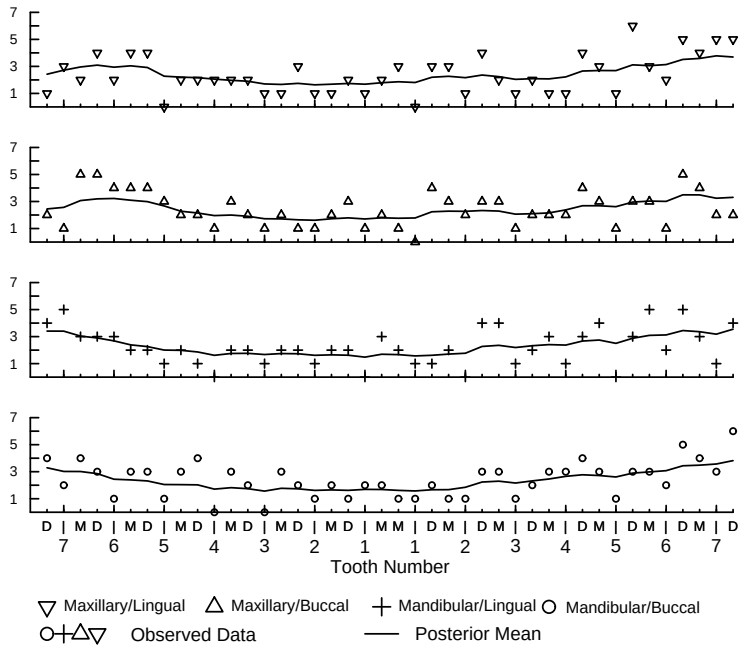
In the MLM for this ICAR, two columns in the RE design matrix $\mathbf{Z}$ describe the difference {avg direct sites} minus {avg interprox sites}, one each for upper, lower jaws.

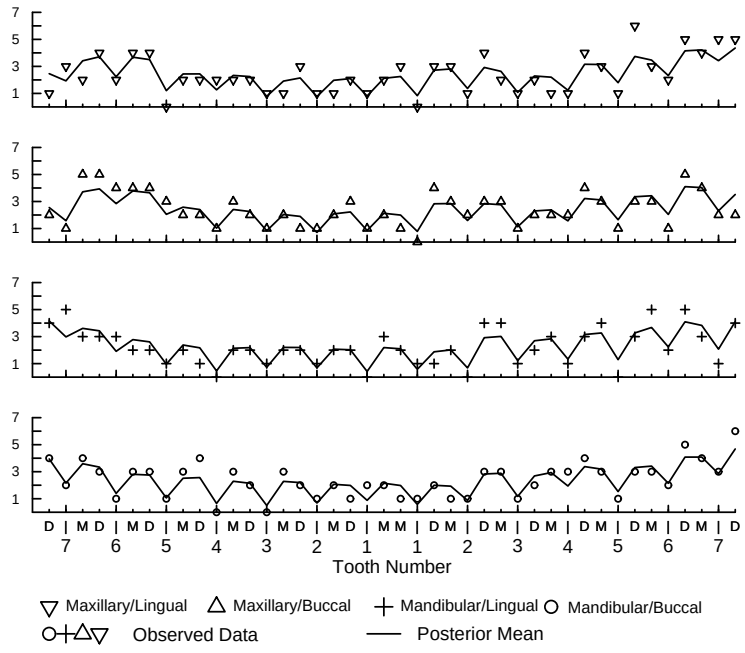
These contrasts are columns of the RE's design matrix, so

- the same contrasts in the data should provide information about σ_s^2
- if we make these two columns fixed effects by moving them from $\mathbf{Z}$ to $\mathbf{X}$ and giving their coefficients flat priors, $\hat{\sigma}_s^2$ should get smaller and the fit should be smoother.

Right?

Wrong: $\hat{\sigma}_s^2$ got bigger and the fit got less smooth



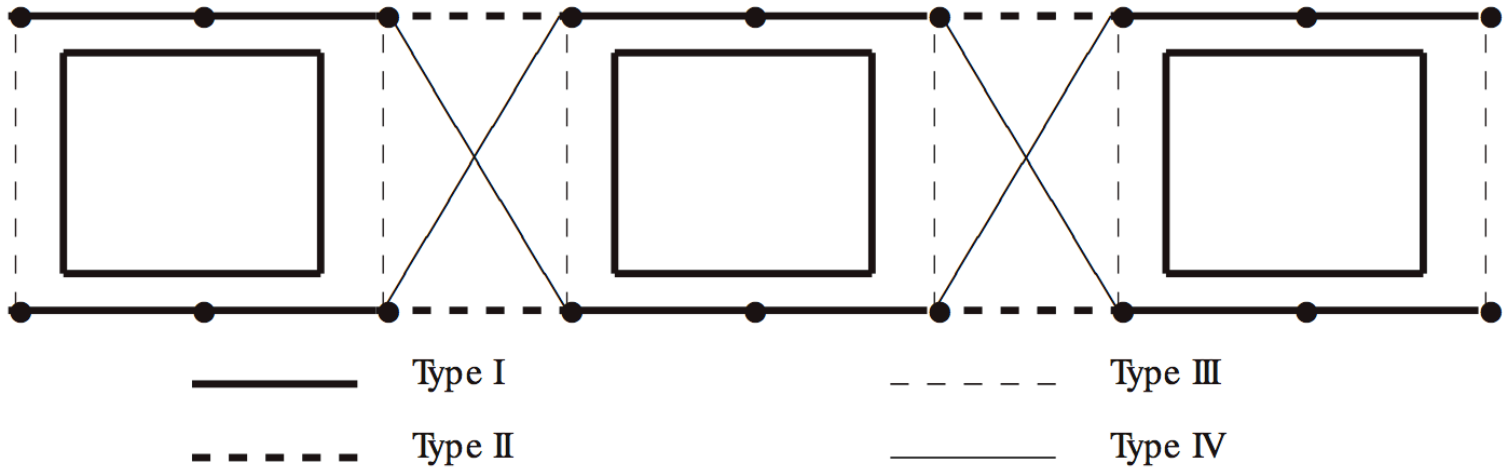


Posterior medians

- σ_s^2: up from 0.25 to 0.41
- σ_e^2: down from 1.25 to 0.63
- ratio σ_e^2/σ_s^2: down from 4.0 to 1.6
- DF in the fit: up from 23.5 to 45.6; only 2 are for the new FEs.

2. Perio data and ICAR with two classes of neighbors

Periodontal neighbor pairs are of four distinct types.



The simple ICAR assumes CAL is equally similar within all neighbor pairs, but it's not.

We considered ICAR-like models with neighbor pairs partitioned into two classes, with different degrees of similarity in the two classes.

- Classification A: side (I, II) vs. interproximal (III, IV)
- Classification B: direct (I) vs. interproximal (II, III, IV).

Two-neighbor-relation CAR (2NRCAR; Reich et al 2007)

$\mathbf{y}$, the n-vector of CAL measurements, is modeled as $\mathbf{y} = \boldsymbol{\delta} + \boldsymbol{\epsilon}$

- $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma_e^2 \mathbf{I}_n)$
- $\boldsymbol{\delta}$ has this improper density

$$f(\boldsymbol{\delta}|\sigma_{s1}^2, \sigma_{s2}^2) \propto \exp\left[-\frac{1}{2}\boldsymbol{\delta}'\left(\mathbf{Q}_1/\sigma_{s1}^2 + \mathbf{Q}_2/\sigma_{s2}^2\right)\boldsymbol{\delta}\right],$$

- σ_{sk}^2 controls the similarity of class k neighbor pairs, $k = 1, 2$
- $\mathbf{Q}_k = (q_{ij,k})$ is $n \times n$ and encodes class-k neighbor pairs
  - $q_{ii,k}$ = number of site i's class-k neighbors
  - $q_{ij,k} = -1$ if sites i and j are class-k neighbors, 0 otherwise.

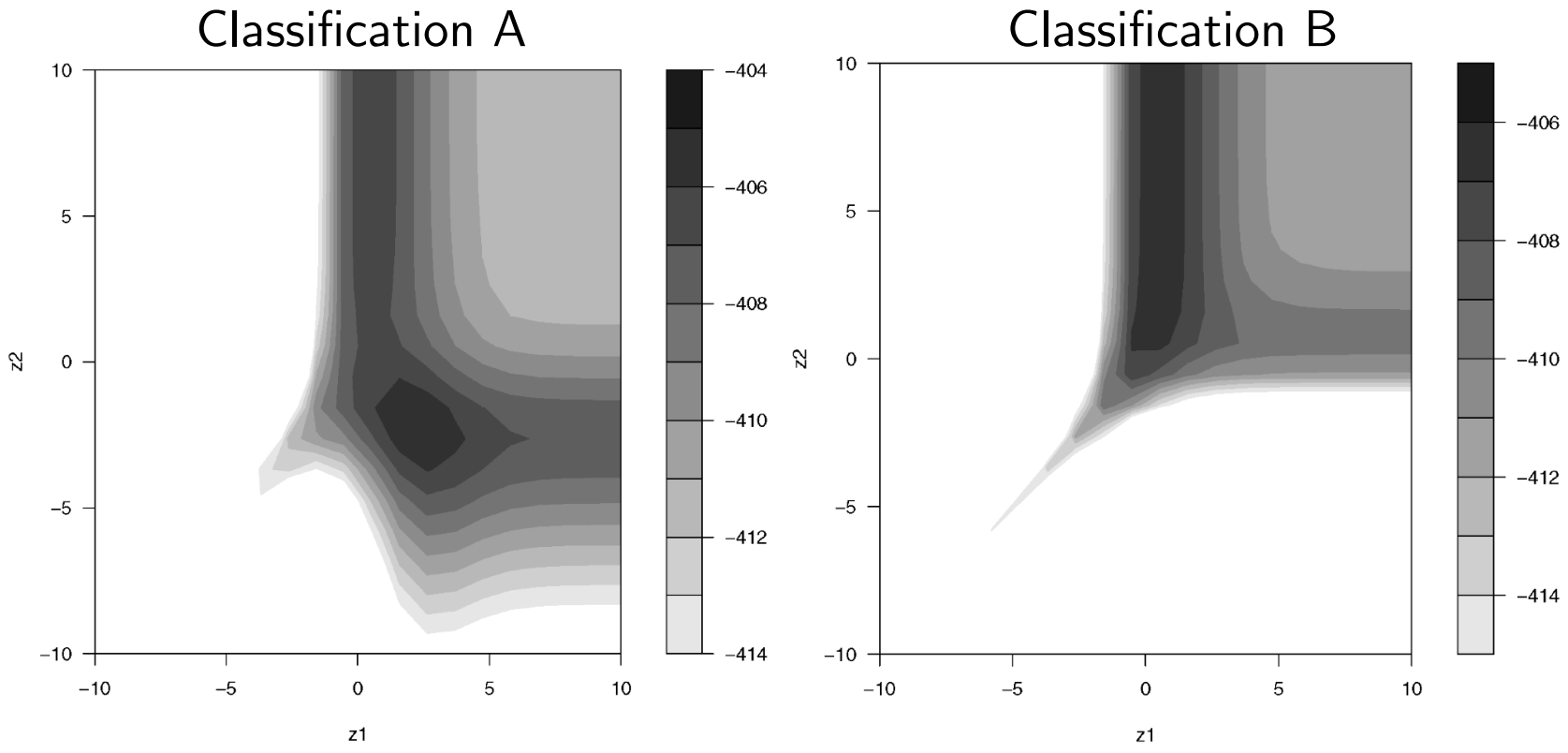
This can be written as a mixed linear model.

Seems simple enough, but the MCMCs had high lagged autocorrelations.

The MCMCs behaved badly because the posterior has a weird shape.

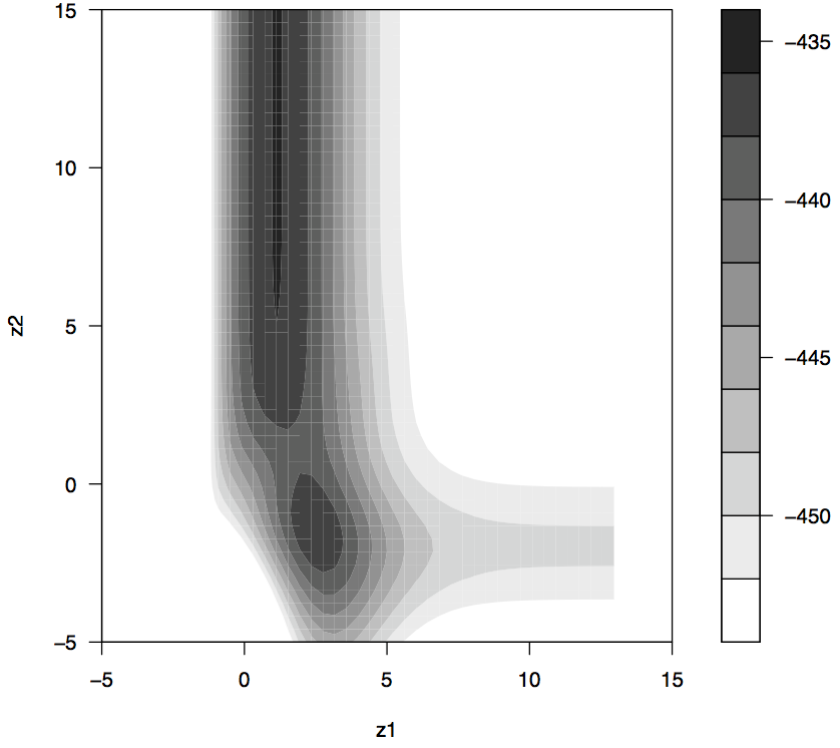
Here are contours of the marginal posterior of (z_1, z_2), $z_k = \log(\sigma_e^2/\sigma_{s1}^2)$.

For one person's data

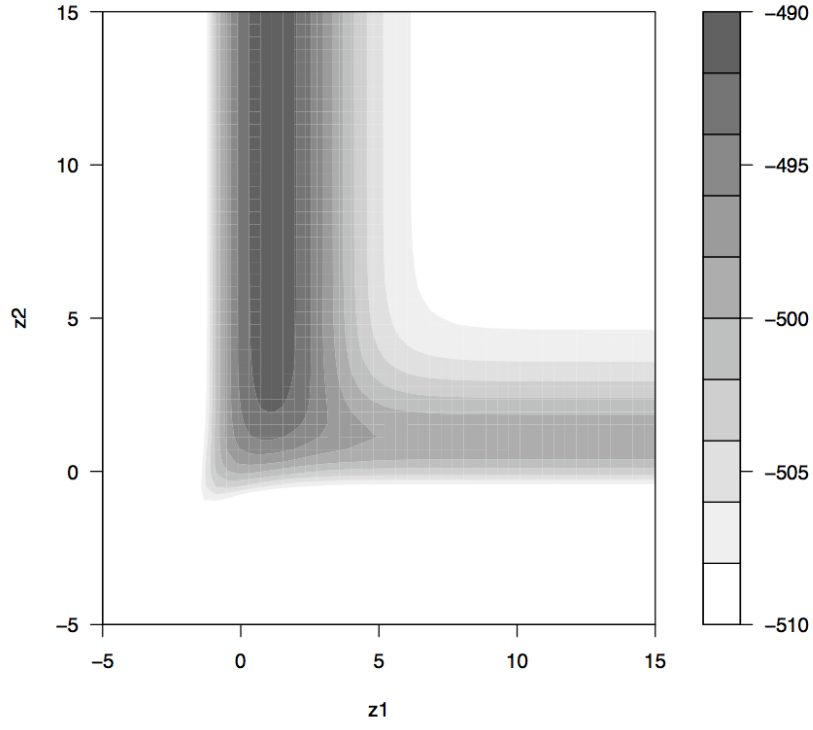


They're even weirder if you analyze data for three people (assuming the same parameter values for all three).

Classification A



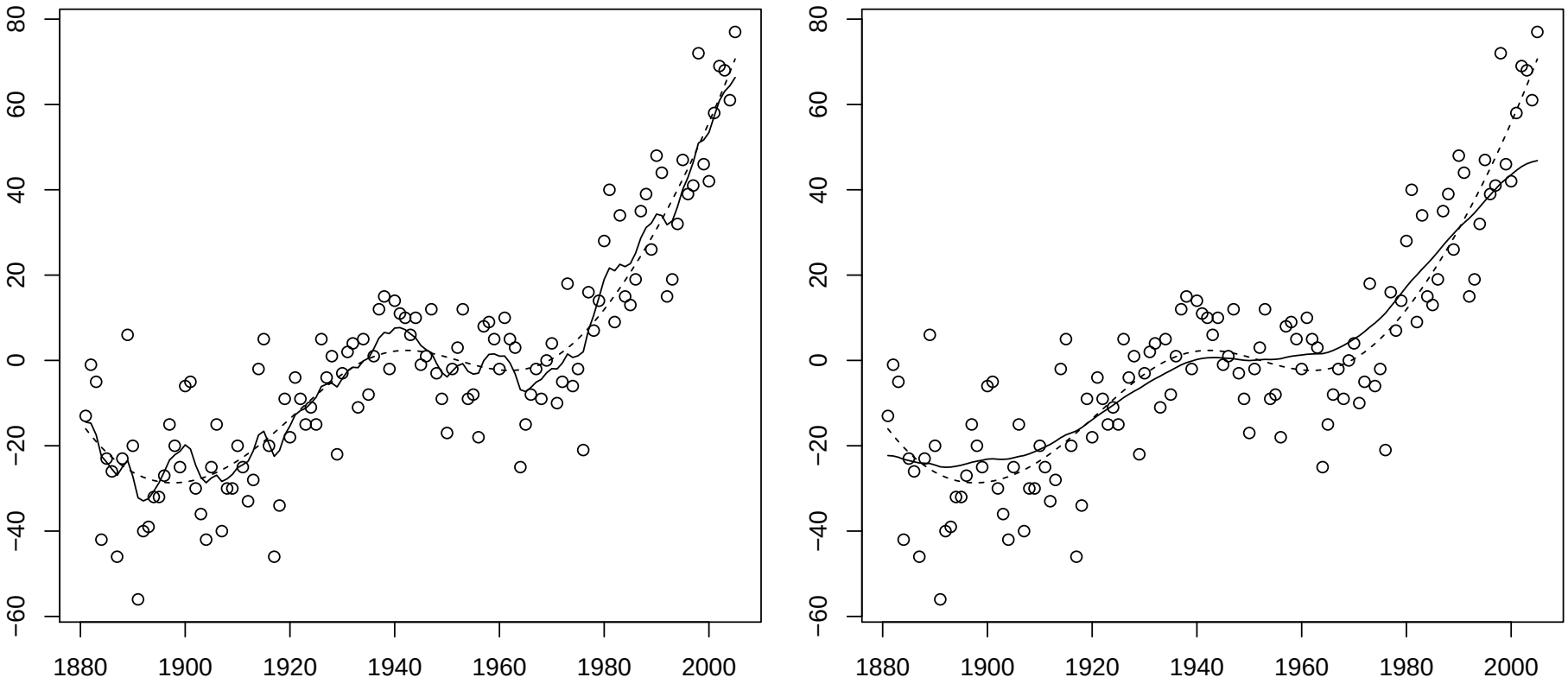
Classification B



3. Very different smooths of the same GMST data

Left: Max RL for penalized spline (dash, 6.7 DF), ICAR (solid, 26.5 DF)

Right: Penalized spline (dash, 6.7 DF); ICAR with 6.7 DF (solid)



If you force the ICAR fit to have 6.7 DF, it allocates them stupidly.

4. Misleading zero variance estimates (Epidermal nerve density, revisited)

The investigators want to compare biopsy (old) vs. blister (new).

19 subjects; at calf and foot, 2 blisters.

They're interested in the between-blister variation.

Here are the max-RL estimates of variance components (CIs from SAS)

Variance component	Variance Estimate	Confidence Interval	
Subject	18,031	8,473	61,169
Subject-by-Location	9,561	4,684	29,197
Blister within Subject/Location	0	0	0
Residual	6,696	5,181	8,992

The CI of 0 to 0 is silly.

The 0 estimate is wrong – blisters do vary.

Zero happens to maximize the RL, but:

- Fix the other three var comps at their RL-maximizing values
- Increase the between-blister variance from zero
- It reaches 1,550 before the log RL declines by 2.

The RL itself supports values much larger than zero.

There's no reason to conclude the between-blister variance is zero.

BUT I had to do a lot more work to establish that.

5. Multiple maxima in the posterior and RL: HMO premium data revisited

The model in Hodges (1998), Sec. 5.2 included state average expenses/hospital admission and a New England indicator.

- flat priors on fixed effects and $1/\sigma_e^2$ [bad idea];
- gamma prior for $1/\sigma_s^2$ with mean 11 and variance 110.

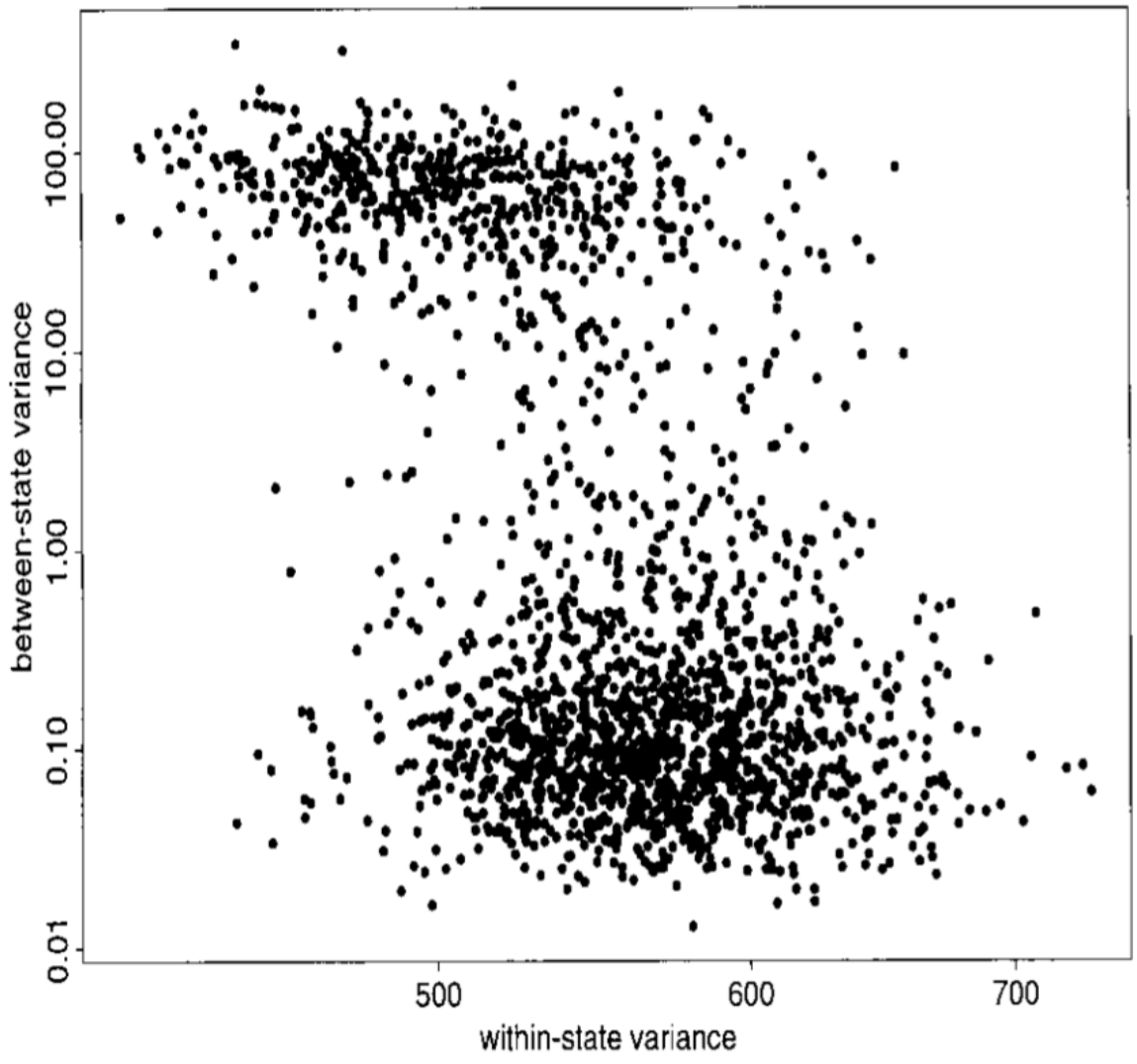
I ran my Gibbs sampler for 1000 iterations [extremely bad idea] and Rao-Blackwellized. I observed:

"The Gibbs sampler was unstable through about 250 iterations; from iteration 251 to 1000 the sampler provided satisfactory convergence diagnostics", so I used those 750 iterations.

I.e., the trace plot for σ_s^2 was stable for the first 250 draws, dropped to a lower level, and then was stable for the last 750 draws.

$\Rightarrow$ Bimodal posterior, but that possibility had never occurred to me.

Wakefield ran *his* Gibbs sampler for 10,000 iterations and got this:



"there are two competing explanations for the observed variability"

The rest of the course . . .

These mysterious, inconvenient, or wrong results can be explained, at least to some extent.

The main tool: Re-expressing the restricted likelihood so all square matrices are diagonalized, so matrix expressions $\rightarrow$ scalar expressions.

A large sub-class of mixed linear models can be re-expressed this way, though not all of them.

This is therefore a place to start, not the last word.

The rest of the course develops and applies this re-expression.