Zero variance estimates

Hardly anything is known about estimates on the boundary of the parameter space.

When a zero variance estimate maximizes the RL, I want to know:

- are the data consistent with "large" values of that variance,
- i.e., does the RL have a flat left tail in that variance?

This section

- examines the BOWREM in some detail;
- gives short comments about a nested ANOVA model; and
- finishes with some thoughts about tools.

Balanced one-way RE model (BOWREM)

$y_{ij} = \beta_0 + u_i + \epsilon_{ij}$, N groups, m per group, $u_i \sim N(0, \sigma_s^2)$, $\epsilon_{ij} \sim N(0, \sigma_e^2)$.

Define $n = Nm$, $S_E = \sum_{ij}(y_{ij} - \bar{y}_{i.})^2$, $S_M = \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$

The log RL is:

$$
\begin{aligned}
&-\frac{n-N}{2}\log(\sigma_e^2) - \frac{1}{2}\frac{S_E}{\sigma_e^2} \\
&-\frac{N-1}{2}\log(\sigma_s^2 m + \sigma_e^2) - \frac{m}{2}\frac{S_M}{(\sigma_s^2 m + \sigma_e^2)},
\end{aligned}
$$

It is maximized by:

$$
\begin{array}{lllll}
\text{if } \frac{S_M}{N-1} & \geq \frac{S_E}{m(n-N)} & \hat{\sigma}_e^2 = & S_E/(n-N) \\
 & & \hat{\sigma}_s^2 = & S_M/(N-1) - \hat{\sigma}_e^2/m \\
 & & & \\
\text{if } \frac{S_M}{N-1} & < \frac{S_E}{m(n-N)} & \hat{\sigma}_e^2 = & (S_E + mS_M)/(n-1) \\
 & & \hat{\sigma}_s^2 = & 0.
\end{array}
$$

When is this RL flat near $\hat{\sigma}_s^2 = 0$?

Consider the log RL's derivative wrt σ_s^2, evaluated at $\sigma_s^2 = 0$ and $\hat{\sigma}_e^2$:

$$\frac{\partial \log \ \mathrm{RL}}{\partial \sigma_s^2} = \left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} \frac{N-1}{2} \left(\frac{S_M}{N-1} \left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} - 1 \right).$$

If $S_M/(N-1) < \hat{\sigma}_e^2/m$, this derivative < 0 and $\hat{\sigma}_s^2 = 0$.

The derivative is small in absolute value if either

- $\hat{\sigma}_e^2/m$ is large or
- $\dfrac{S_M}{N-1} \left(\dfrac{\hat{\sigma}_e^2}{m}\right)^{-1}$ is close to 1.

These two conditions have different implications.

$$\frac{\partial \log \text{ RL}}{\partial \sigma_s^2} = \left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} \frac{N-1}{2} \left(\frac{S_M}{N-1}\left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} - 1\right).$$

<u>If $\hat{\sigma}_e^2/m$ is large</u>, the design and data provide low resolution for σ_s^2.

- A wide interval of positive σ_s^2 have RL near the max value.
- The RL provides this information; it is routinely ignored.

To increase the design's resolution and get $\hat{\sigma}_s^2 > 0$:

- Increase m, holding constant N, S_M, and $\hat{\sigma}_e^2$.
- Simply increasing N doesn't work:
  - Holding constant $S_M/(N-1)$ and $\hat{\sigma}_e^2/m$, this leaves the key condition $S_M/(N-1) < \hat{\sigma}_e^2/m$ unchanged.
  - The derivative does become larger in magnitude (steeper dropoff).

$$\frac{\partial \log\ \mathrm{RL}}{\partial \sigma_s^2} = \left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} \frac{N-1}{2} \left(\frac{S_M}{N-1} \left(\frac{\hat{\sigma}_e^2}{m}\right)^{-1} - 1\right).$$

If $\hat{\sigma}_e^2/m - S_M/(N-1) < 0$ but close to 0,

$\partial \log \mathrm{RL}/\partial\sigma_s^2 < 0$ and small because the peak is close to $\sigma_s^2 = 0$.

The RL may but does not necessarily decline slowly from $\sigma_s^2 = 0$

- When $S_M/(N-1) > 0.5\hat{\sigma}_e^2/m$, $\partial^2 \log \mathrm{RL}/\partial(\sigma_s^2)^2 < 0$
- and the restricted likelihood can drop off quickly.

Otherwise, it drops off slowly.

A more complicated ANOVA (Epidermal nerve density)

The real example:

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

Consider a simplified version with 20 subjects and balance

Location is a fixed effect

Four variance components:

- σ^2_{s1} variation between subjects
- σ^2_{s2} variation between subjects in the difference between location
- σ^2_{s3} variation between blisters within subject and location
- σ^2_e ($j = 4$) error (variation between images within a blister)

$$\log RL(\sigma^2_{s1}, \sigma^2_{s2}, \sigma^2_{s3}, \sigma^2_e | y) = -0.5 \sum_{j=1}^{4} DF_j \left[\log \theta_j + \hat{\theta}^u_j / \theta_j \right],$$

DF_j are the usual ANOVA DF; $\hat{\theta}^u_j$ are the usual mean squares.

Degrade the design's resolution by increasing the error mean square, $\hat{\theta}^u_4$. What happens?

Mean squares are 15, 7, and 3 for sub, sub$\times$loc, blister(sub$\times$loc)

$\hat{\theta}_4^u$	$\hat{\sigma}_{s1}^2$	$\hat{\sigma}_{s2}^2$	$\hat{\sigma}_{s3}^2$	$\hat{\sigma}_e^2$
1	1.00	1.00	1.00	1.00
2	1.00	1.00	0.50	2.00
3	1.00	1.00	0.00	3.00
4	1.00	0.83	0.00	3.67
...		...		...
8	1.00	0.17	0.00	6.33
9	1.00	0.00	0.00	7.00
10	0.93	0.00	0.00	7.58
11	0.86	0.00	0.00	8.15
...		...		...
21	0.14	0.00	0.00	13.91
22	0.06	0.00	0.00	14.48
23	0.00	0.00	0.00	15.05

The analogous thing happens if we fix error MS and increase $\hat{\theta}_3^u$.

This is unexplained (as far as I know).

Some thoughts about tools

My question: When a zero variance estimate maximizes the RL, are the data consistent with "large" values of that variance?

The obvious form for this information is a simple one-sided CI.

- We might use derivatives; but how to calibrate "small"?
- A simple CI, if one exists, avoids this calibration problem.
- This would be useful to Bayesians because it's simple and fast.
- And it's OK if the CI's coverage tends to be low.

The obvious candidate would use the profile log RL.

- Upper end: $\sigma_s^2 \ni$ profile log RL is reduced by c from its max.
- Some software already computes the profile RL.
- Problem: Which c to use? Solution: Big simulation experiment.