Frequentist Analysis Strategy for Complex Aliasing

- Some properties of Method II (cont.)
  - The constraint imposed by effect heredity substantially reduces the model space searched in Method II.
    - Consider, the 12-run PBD with 11 factors and h=4 in Method II.
    - The total number of main effects and two-factor interactions is $11+\binom{11}{2}=66$.
    - The total number of models with five terms (and the intercept) is $\binom{66}{4}=720,720$.
    - There are 15,510 models satisfying the effect heredity requirement. This is about 2.2% of the total number of models.
  - As the number of factors and h increase, the search will become computationally prohibitive. An alternative is to use an efficient stochastic search such as the *Bayesian variable selection.*

❖ **Reading**: textbook, 9.4

Analysis of Cast Fatigue Experiment

- The design is an $OA(12,2^7)$.
- Consider the main-effect-only model.
  - The 11 columns in Table 1 are mutually orthogonal.
  - The main effect estimates are uncorrelated and unbiased estimates of the main effects *if there are no interactions*.
  - Half-normal plot:

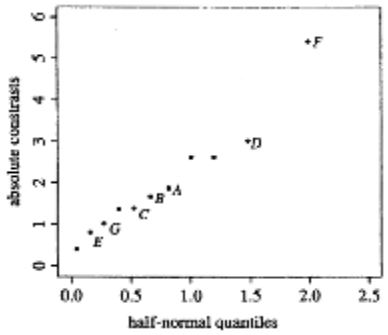

Figure 1: Half-normal plot, cast fatigure experiment

  - The model with F alone has an $R^2=0.45$.
  - The model (F,D) has an $R^2=0.59$, and the fitted model is

$$\hat{y}=5.73+0.458F-0.258D.$$

Analysis of Cast Fatigue Experiment

- Entertain a model with ME F and all the interactions involving F.
  - Identify a significant FG interaction.
  - The model (F, FG) has an $R^2 = 0.89$
  - The model (D, F, FG) has an $R^2 = 0.92 \Rightarrow D$ appears not significant
  - The model for predicted fatigue lifetimes is

$$\hat{y} = 5.7 + 0.458F - 0.459FG,$$

    - Set $(F, G) = (+, -)$. The predicted life is

$$5.7 + 0.458 - 0.459(-1) = 5.7 + 0.92 = 6.62.$$

    - Compared to the average lifetimes (=5.7), it has a 16% increase (=0.92/5.7)
    - Compared to $(F, G) = (-, -)$, it has a 38% increase (=1.84/4.78)
  - The potential for a dramatic improvement would not have been possible without discovering the FG interaction.

Analysis of Cast Fatigue Experiment

- Using Method I, the model (F, FG) is found in Step 1 and does not change in further steps.
- Method II also identifies model (F, FG) using $h = 2$.
- For larger h (say, 3 or 4), Method II identifies multiple good fitting models, *each of which contains F and FG.*
- The original analysis used main-effect-only model and
  - identified the F and D as significant,
  - noted a discrepancy between the findings and previous work, i.e., the sign of D effect was reversed,
  - concluded that the possible cause was an interaction DE,
  - claimed that the design did not generate enough information to determine DE.
- Because the DE interaction is orthogonal to both D and E main effects, these three effect estimates are uncorrelated and therefore the DE interaction does not affect the sign of the factor D main effect.

Analysis of Cast Fatigue Experiment

- The design's aliasing patterns can be used to explain this apparent reversal.

Table 9: Estimates and Alias Patterns, Cast Fatigue Experiment

Effect	Estimated Effect	Alias Pattern
A	0.326	$A-\frac{1}{3}FG$
B	0.294	$B-\frac{1}{3}FG$
C	−0.246	$C+\frac{1}{3}FG$
D	−0.516	$D+\frac{1}{3}FG$
E	0.150	$E-\frac{1}{3}FG$
F	0.915	F
G	0.183	G
8	0.446	$-\frac{1}{3}FG$
9	0.453	$-\frac{1}{3}FG$
10	0.081	$-\frac{1}{3}FG$
11	−0.242	$+\frac{1}{3}FG$

Analysis of Cast Fatigue Experiment

- The D estimate can be adjusted as follows:

$$\hat{D}-\frac{1}{3}\widehat{FG}=-0.516-(-0.306)=-0.210.$$

- A 95% confidence interval for D, (−0.526, 0.106)=(−0.210 ± 0.316), shows that a positive D main effect is possible.
- The magnitude of the estimated effects of A, B and C can be explained from Table 9 by an FG interaction; i.e.,

$$(-\frac{1}{3}\widehat{FG},-\frac{1}{3}\widehat{FG},\frac{1}{3}\widehat{FG})=(0.306,0.306,-0.306)$$

are close to their estimates, 0.326, 0.294 and −0.246, respectively.

- The effects A-E and 8-11 have the same sign as their FG aliases, which lend further support to the existence of a significant FG interaction.

Bayesian Analysis Strategy for Complex Aliasing

- The Methods I and II work well if
  - the models under search are restricted to have a small number of terms, and
  - the number of candidated terms to choose from is moderate.
- As the number of candidate terms to choose from and/or the number of terms allowed in the model increases,
  - the search over the model space is too incomplete for Method I to be effective, and
  - the required computations for Method II become prohibitively expensive even with the help of effect heredity to reduce the number of models.

Bayesian Analysis Strategy for Complex Aliasing

- Bayesian framework:
  - Let $\mathbf{y}$ denote the data, and θ denote the the vector of parameters. Both are random variables.
  - $f(\mathbf{y}|\theta)$: the likelihood of $\mathbf{y}$ given θ
  - $\pi(\theta)$: the prior distribution for θ.
  - The posterior distribution for θ is:

$$\pi(\theta|\mathbf{y}) = f(\mathbf{y}|\theta)\pi(\theta) / \int f(\mathbf{y}|\theta)\pi(\theta)d\theta$$

  - Using the posterior distribution, inference about θ can then be made.
- Parameters $\theta = (\beta, \delta, \sigma^2)$
  - β: a $(k+1)$ vector containing the intercept and factorial effects,
  - δ: a $(k+1)$ vector of 0's and 1's indicating the significance of the effects
  - σ^2: error variance in the linear model

Bayesian Analysis Strategy for Complex Aliasing

- Assignment of distributions
  - $\mathbf{y}|\beta,\delta,\sigma^2$: a linear model structure

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon,$$

where $\mathbf{X}$ is the $N \times (k+1)$ model matrix, and $\varepsilon \sim MN(\mathbf{0}, \sigma^2 \mathbf{I}_{N\times N})$.

  - $\beta|\delta,\sigma^2$: normal mixture prior

$$\pi(\beta_i|\delta_i,\sigma^2) = \begin{cases} N(0,\sigma^2\tau_i^2) & \text{if } \delta_i = 0 \\ N(0,\sigma^2(c_i\tau_i)^2) & \text{if } \delta_i = 1 \end{cases}.$$

    - When $\delta_i = 0$, the constant τ_i needs to be specified so that β_i is tightly centered around 0 and therefore does not have a large effect.
    - The constant c_i needs to be chosen with $c_i \gg 1$ to indicate the possibility of a large β_i when $\delta_i = 1$.
    - The constants τ_i and c_i should be chosen to represent respectively a "small" effect, and how many times larger a "large" effect should be.

Bayesian Analysis Strategy for Complex Aliasing

- Assignment of distributions (cont.)
  - The δ and σ^2 are independent.
  - δ
    - Independence prior: $\pi(\delta) = \prod_{i=1}^{p+1} p_i^{\delta_i}(1-p_i)^{1-\delta_i}$, where $p_i = Prob(\delta_i = 1)$.
    - Prior with effect heredity principle
      - Consider the example $\delta = (\delta_A, \delta_B, \delta_C, \delta_{AB}, \delta_{AC}, \delta_{BC})$
      .

$$\begin{aligned} Prob(\delta) = \; & Prob(\delta_A)Prob(\delta_B)Prob(\delta_C) \\ & \times Prob(\delta_{AB}|\delta_A,\delta_B)Prob(\delta_{AC}|\delta_A,\delta_C)Prob(\delta_{BC}|\delta_B,\delta_C). \end{aligned}$$

      .

$$Prob(\delta_{AB} = 1|\delta_A,\delta_B) = \begin{cases} p_{00} & \text{if } (\delta_A,\delta_B) = (0,0) \\ p_{01} & \text{if } (\delta_A,\delta_B) = (0,1) \\ p_{10} & \text{if } (\delta_A,\delta_B) = (1,0) \\ p_{11} & \text{if } (\delta_A,\delta_B) = (1,1) \end{cases}.$$

Bayesian Analysis Strategy for Complex Aliasing

- Assignment of distributions (cont.)
  - δ (cont.)
    - Some notes for $Prob(\delta)$:
      - Independence is assumed among the δ_i's for main effects
      - Conditional independence principle: conditional on δ_i's for main effects, the δ_i's for interactions are independent
      - Inheritance principle: the significance of a term depends only on those terms from which it was formed
    - Some notes for $Prob(\delta_{AB})$:
      - We should choose p_{00} small (e.g., 0.01), p_{01} and p_{10} larger (e.g., 0.10) and p_{11} largest (e.g., 0.25) $\Rightarrow$ **relaxed weak heredity**
      - **Strict weak heredity**: setting $p_{00} = 0$
      - **Strong heredity**: setting p_{00}=p_{01}=p_{10}=0
      - Probabilities of less than 0.5 for both main effects and interactions represent the belief that relatively few terms are active, i.e., *effect sparsity* holds.

Bayesian Analysis Strategy for Complex Aliasing

- Assignment of distributions (cont.)
  - σ^2: inverse gamma

$$\sigma^2 \sim \text{IG}(\nu/2, \nu\lambda/2),$$

whose density is

$$\pi(\sigma^2) \propto (\sigma^2)^{-(\nu/2+1/2)} \exp\{-\nu\lambda/(2\sigma^2)\}.$$

- The evaluation of the posterior for θ can be implemented by using **Gibbs sampling**, a simple *Markov Chain Monte Carlo* (MCMC) technique for drawing samples from a posterior distribution (read textbook, 9.5.2, for details).
- Because δ specifies a model by the i's with $\delta_i = 1$, the posterior for δ is of particular interest.
- Choice of prior tuning constants τ, **c**, ν, λ: see textbook, 9.5.3, for details.

❖ **Reading**: textbook, 9.5