Hierarchical modeling of data gathered (1)
with (single-stage) cluster sampling:
an example of how non-Bayesian
estimates can produce rubbish results

(many other (quite different) examples can be produced)

Absurd unbiased estimator

I'm studying quality of hospital care in California in 2011; to do this I take a (single-stage cluster sample) random sample of I hospitals (in CA in 2011) and a random sample of J patients from the chosen hospitals (should really have the flexibility to take J_i patients in hospital i, but let's look here at $J_i = J \ \forall i$), and I measure a real-valued quality-of-care (QoC) score y_{ij} for patient j in hospital i.

To really understand what's going on I should also measure hospital-level and patient-level predictor variables (covariates) that help to explain QoC differences, but even if I had these variables, the following <u>variance-components</u> model would be a good starting point:

$$Y_{ij} = \mu + a_i^H + a_{ij}^P$$

$a_i^H \overset{\text{iid}}{\sim} N(0, \sigma_H^2)$, indep, $a_{ij}^P \overset{\text{iid}}{\sim} N(0, \sigma_P^2)$

$\left(\begin{array}{l} i = 1, \dots, I \leftarrow \text{hosp.} \\ j = 1, \dots, J \leftarrow \text{pt.} \end{array} \right.$

the complete parameter vector in this model is $(\mu, \sigma_H^2, \sigma_P^2) = \eta$

and interest typically focuses on the <u>grand mean</u> μ, the <u>variance</u> <u>components</u> σ_H^2 and σ_P^2, and functions of σ_H^2 and σ_P^2 such as $\dfrac{\sigma_H^2}{\sigma_H^2 + \sigma_P^2}$

($\mu \leftarrow$ <u>fixed</u> effect is the population mean quality of care score in CA in 2011; a_i^H (<u>random</u> effect) is the deviation of mean quality at hospital i from the grand mean; a_{ij}^P (<u>random</u> effect) is the deviation of quality for patient j from the mean quality at hospital i; σ_H^2 ← (variance component) quantifies variations from hospital to hospital (between hospitals) in mean quality; σ_P^2 ← (variance component) quantifies variations within hospitals (between patients) in patient-level quality,

$$V_{RS}(Y_{ij}) = V_{RS}(\mu + a_i^H + a_{ij}^P) = \sigma_H^2 + \sigma_P^2$$

is the total variance in quality; and

$\left(\frac{\sigma_H^2}{\sigma_H^2 + \sigma_P^2} \right)$ is the proportion of the total variance "accounted for" by variations in mean hospital quality. Following Fisher (1932), a standard frequentist analysis of this model is based on the analysis of variance table on the next page; you're supposed to compute sums of squares, degrees of freedom, and mean squares.

source	sums of squares	degrees of freedom	mean squares
between hospitals	$SS_H = J \sum_{i=1}^{I} (y_{i\cdot} - y_{\cdot\cdot})^2$	$(I-1)$	$MS_H = \frac{SS_H}{I-1}$
within hospitals	$SS_P = \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - y_{i\cdot})^2$	$IJ - I = I(J-1)$	$MS_P = \frac{SS_P}{I(J-1)}$
total	$SS_T = \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - y_{\cdot\cdot})^2$	$IJ - 1$	

here the dot subscripts denote averaging:

$$y_{i\cdot} = \frac{1}{J} \sum_{j=1}^{J} y_{ij} \quad \text{and} \quad y_{\cdot\cdot} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij} .$$

~~Now focus on the variance components;~~

you can show the following things:

$$\boxed{E_{RS}(MS_P) = \sigma_P^2 \quad \text{and} \quad E_{RS}(MS_H) = J\sigma_H^2 + \sigma_P^2}$$

so if you're a fan of unbiasedness you would

adopt a kind of method-of-moments attitude and equate

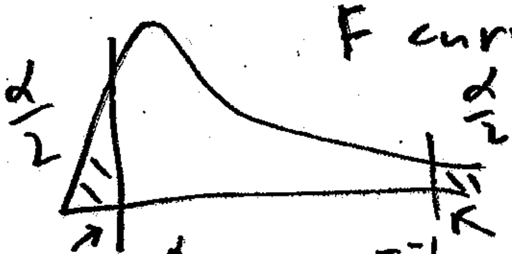
$$\begin{Bmatrix} MS_p = \hat{\sigma}_p^2 \\ MS_H = J\hat{\sigma}_H^2 + \hat{\sigma}_p^2 \end{Bmatrix}$$ and solve to get (5)

$$\begin{Bmatrix} \hat{\sigma}_p^2 = MS_p \\ \hat{\sigma}_H^2 = \dfrac{MS_H - MS_p}{J} \end{Bmatrix}.$$

These are definitely unbiased estimates of the variance components, but $\hat{\sigma}_H^2$ is absurd because it can easily go negative. This doesn't bother Scheffé (1959) (pp. 228-229), but he reveals with his attitude that he's never done any serious applied work. He goes on to derive the following confidence interval for $\theta = \frac{\sigma_H^2}{\sigma_p^2}$, which is related monotonically to $\frac{\sigma_H^2}{\sigma_H^2 + \sigma_p^2}$

$$= \frac{1}{1 + \frac{\sigma_p^2}{\sigma_H^2}} = \frac{1}{1 + \frac{1}{\theta}} = \frac{\theta}{\theta + 1}:$$

$$\frac{1}{J}\left(\frac{MS_H/MS_P}{F^{-1}_{I-1,\,I(J-1)}\left(1-\frac{\alpha}{2}\right)}-1\right) < \underset{\frac{\sigma^2_H}{\sigma^2_P}}{\underset{\uparrow}{\theta}} < \frac{1}{J}\left(\frac{MS_H/MS_P}{F^{-1}_{I-1,\,I(J-1)}\left(\frac{\alpha}{2}\right)}-1\right)$$



F curve with $(I-1)$ and $I(J-1)$ degrees of freedom

$\frac{\alpha}{2}$ $\qquad$ $\frac{\alpha}{2}$

$F^{-1}_{I-1,\,I(J-1)}\left(\frac{\alpha}{2}\right)$ $\qquad$ $F^{-1}_{I-1,\,I(J-1)}\left(1-\frac{\alpha}{2}\right)$

Not only is it possible for this interval to have a negative left endpoint; the entire interval for $\left(\frac{\sigma^2_H}{\sigma^2_P}\right)$ can be negative! This also doesn't bother Scheffé (1959, pp. 229-231); his position on this issue shows that he values <u>unbiasedness</u> more highly than <u>sanity</u>. (R code & winbugs demo in class)