SRS with Repeated Measure of a Subject and Response Error

Background and Setting: The Population

Label the subjects in the population by \( s = 1, ..., N \), and represent the response for subject \( s \) as \( y_s \) and let a parameter for each subject be represented by \( \mu = \begin{pmatrix} \mu_1 & \mu_2 & \cdots & \mu_N \end{pmatrix}' \). Define the population mean as

\[
\mu = \frac{1}{N} \sum_{s=1}^{N} \mu_s \quad \text{and define the population variance as} \quad \sigma^2 = \frac{1}{N-1} \sum_{s=1}^{N} (\mu_s - \mu)^2 \quad \text{where}
\]

\[
\sigma^2 = \frac{1}{N-1} Y' (I_N - J_N / N) Y = \frac{1}{N-1} \sum_{s=1}^{N} (\mu_s - \mu)^2.
\]

We can represent the parameter for subject \( s \) is by

\[
\mu_s = \mu + \beta_s \quad (1)
\]

where \( \beta_s = \mu_s - \mu \). This model is not stochastic. Also, using the definition of \( \mu \), we see that \( \sum_{s=1}^{N} \beta_s = 0 \).

Sampling Subjects

Suppose we have a simple random (without replacement) sample of \( n \) subjects. From previous developments, note that we can represent a random permutation of the subjects by the vector \( Y = ((Y_i)) \), then \( E_p(Y) = I_n \mu \) and \( \text{var}_p(Y) = \sigma^2 \left(I_n - \frac{J_n}{N}\right) \). The subscript \( p \) indicates expectation with respect to sampling design probabilities. For a simple random sample of \( n \) subjects, \( E_p(\bar{Y}_i) = 1_n \mu \) and \( \text{var}_p(\bar{Y}_i) = \sigma^2 \left(I_n - \frac{J_n}{N}\right) \).

Suppose now that we represent a vector where we repeat each random variable \( p \) times. We can represent this new vector as \( Y^* = \bar{Y}_i \otimes 1_p = \left(I_n \otimes 1_p\right) \bar{Y}_i = C' \bar{Y}_i \), where \( C' = \left(I_n \otimes 1_p\right) \). Using properties of expected values and variances, \( E_p(C' \bar{Y}_i) = C' 1_n \mu \) and

\[
\text{var}_p(C' \bar{Y}_i) = \sigma^2 C' \left(I_n - \frac{J_n}{N}\right) C. \quad \text{These expressions can be simplified using the result that} \quad C' 1_n = \left(I_n \otimes 1_p\right) 1_n = 1_{np}, \quad \text{and}
\]
In summary, when \( Y_i' \) = \((Y_1' \ Y_2' \ \cdots \ Y_n')' \), where \( Y_i = ((Y_i)) = (Y_i \ Y_i \ \cdots \ Y_i)' \),

\[ E_p \left( Y_i' \right) = 1_p \mu \text{ and } \text{var}_p \left( Y_i' \right) = \sigma^2 \left( I_n - \frac{J_n}{N} \right) \otimes J_p. \]

SRS with Response Error Model

Additions to Background and Setting: The Response Error Model for Subjects in the Population

In the population, we assume that on a subject we observe a response which may have a different value than the subject’s parameter. This difference we call response error. We represent the \( k^{th} \) response for subject \( s \) as \( Y_{sk} \). We assume the expected value over response error \( E_r(Y_{sk}) = \mu_s \). A response error model for \( k^{th} \) response for subject \( s \) is given by

\[ Y_{sk} = \mu_s + E_{sk}. \]

Let us represent the \( \text{var}_p \left( Y_{sk} \right) = \sigma^2_s \). Also, let us assume that multiple measures of response on a subject are independent. We represent response for \( p \) measures on subject \( s \) as \( Y_s = \left( Y_{s1} \ Y_{s2} \ \cdots \ Y_{sp} \right)' \), and find that with these assumptions, \( E_r(Y_s) = 1_p \mu_s \) and \( \text{var}_p \left( Y_s \right) = \sigma^2_s 1_p \). Then,

\[
\text{var}_r \left( \begin{array}{c} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{array} \right) = \left( \begin{array}{cccc} \sigma^2_1 1_p & 0_p & \cdots & 0_p \\ 0_p & \sigma^2_2 1_p & \cdots & 0_p \\ \vdots & \vdots & \ddots & \vdots \\ 0_p & 0_p & \cdots & \sigma^2_N 1_p \end{array} \right).
\]

Simple random sampling and response error- a mixed model.

We combine the ideas of response error on subjects, and sampling of subjects. In so doing, we assume that the processes of sampling and response error are independent. We represent response for the \( k^{th} \) measure on the \( i^{th} \) selected subject as \( Y_{ik} \). Let \( B_i \) represent the deviation that corresponds to the \( i^{th} \) selected subject’s expected response, from the population mean. This term is
random, and varies depending upon which subject is selected on the $i^{th}$ draw. The mixed model is given by

$$Y_{ik} = \mu + B_i + E_{ik}.$$ 

Suppose that a simple random sample of $n$ subjects are selected, with $p$ measures of response made on each selected subject. We represent this response vector as $Y_i = (Y_{i1}' \ Y_{i2}' \cdots \ Y_{ip}')'$, where $Y_i = ((Y_{ik})) = (Y_{i1} \ Y_{i2} \cdots \ Y_{ip})'$. Now, $E_{pr}(Y_{ik}) = E_p[E_{ri,p}(Y_{ik})]$. Given a selected sample, the $i^{th}$ selected subject is fixed (and known), so that $E_{ri,p}(Y_{ik}) = E_{ri,p}(Y_{i,k} | i^{th} selection is subject $s_i$) = $\mu_{s_i} = Y_i$.

Then $E_{pr}(Y_{ik}) = E_{p}[Y_i] = \mu$. As a result, $E_{pr}(Y_i) = I_{np}\mu$.

Next, we evaluate $\text{var}_{pr}(Y_i) = \text{var}_p[E_{ri,p}(Y_i)] + E_p[\text{var}_{ri,p}(Y_i)]$. Now $E_{ri,p}(Y_i) = Y_i^*$, where $E_p(Y_i^*) = I_{np}\mu$ and $\text{var}_p(Y_i^*) = \sigma^2\left(I_n - \frac{J}{N}\right) \otimes J_p$. Also, \[
\text{var}_{ri,p}(Y_i) = \text{var}_{ri,p}\left(\begin{array}{c}
Y_{s_1} \\
Y_{s_2} \\
\vdots \\
Y_{s_n}
\end{array}\right) = \sigma^2 I_p \begin{bmatrix}
0 & 0 & \cdots & 0 \\
0 & \sigma^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma^2
\end{bmatrix}.
\]

Taking the expected value over all possible samples, $E_p[\text{var}_{ri,p}(Y_i)] = \sigma^2 \left(I_p \otimes I_p\right)$. Combining these terms, \[
\text{var}_{pr}(Y_i) = \sigma^2 \left(I_n - \frac{J}{N}\right) \otimes J_p + \sigma^2 \left(I_n \otimes I_p\right).
\]

We will often assume that the population size is large relative to the sample size, and thus drop the term in the expression for the variance that is divided by $N$. With this assumption, the mixed model is given by

$$Y_{ik} = \mu + B_i + E_{ik}$$

or

$$Y_i = X_i\mu + Z_iB + E_i$$

or

$$Y = X\mu + ZB + E.$$ 

where $X = 1_{np}$, $Z = 1_n \otimes 1_p$, $E_{pR}(Y_i) = 1_{np} \mu$, and $\text{var}_{pR}(Y_i) = 1_n \otimes (J_p \sigma^2 + I_p \sigma_e^2)$.

Note that $\text{var}_{pR}(Y_i) = J_p \sigma^2 + I_p \sigma_e^2 = \begin{pmatrix}
(\sigma^2 + \sigma_e^2) & \sigma^2 & \cdots & \sigma^2 \\
\sigma^2 & (\sigma^2 + \sigma_e^2) & \cdots & \sigma^2 \\
\sigma^2 & \sigma^2 & \cdots & (\sigma^2 + \sigma_e^2) \\
\vdots & \vdots & \ddots & \vdots \\
\sigma^2 & \sigma^2 & \cdots & (\sigma^2 + \sigma_e^2)
\end{pmatrix}$. This matrix is compound symmetric.