Basic Ideas and the Random Permutation (RP) model for sampling from a finite population.

Ed Stanek

1. FINITE POPUTION- BASICS

Frames
Labels, Response
Notation and terminology
Response: y_s
Potentially Observable Response (Experiments): y_{st}
Parameters:

$$\mu = \frac{1}{N}\sum_{s=1}^{N} y_s\,,\quad \beta_s = y_s - \mu\,,\quad \sigma^2 = \frac{1}{N-1}\sum_{s=1}^{N}\left(y_s - \mu\right)^2$$

Notation and terminology (Potentially Observable)
Expansions
Multi-level (clusters, units)
Measurement Error (replication)
Treatments (conditions)
Multiple Variables (Relationships)
Missing Response

2. GENERAL PROBLEM OF INFERENCE

Objective: Guess a target parameter or realized random variable (or a set) in a population.

<u>Design Based</u>
Observe: Response for part of the population (a sample)
Strategy:
Use assumptions about the sample to define the 1st and 2nd moments of random variables for the population.
Express the 'target' as a linear function of sample and other terms.
Define a predictor (of the other terms) as a linear function of the sample
Require the predictor to be unbiased (unconditionally)
Require the predictor to have minimum MSE (unconditionally)

<u>Model Based</u>
Strategy:
Assume a distribution of random variables (or moments of a superpopulation).
Determine properties of estimators.
Derive optimal estimators
Observe data (realizations of random variables in estimator)

Bayesian Framework

Prior distribution- distribution (i.e. define 1st and 2nd moments, and likelihood) of the random variables)

Observed data (define the sample observed data)

Posterior distribution- develop the posterior distribution conditional on the sample.

3. REVIEW OF BASIC STATISTICS

Constant: lower case Latin letters, i.e. y_s where s is a label identifying subject

Labels: identifier for a physical unit in a population

Random variable: upper case Latin letters, i.e. Y_i where i identifies the position of a random variable in a sequence (or a set if exchangeable).

Expected Value (notation): $E(Y_i)=\sum_s P(Y_i = y_s)y_s$

Variance (notation): $\text{var}(Y_i)=\sum_s P(Y_i = y_s)\left(y_s - E(Y_i)\right)^2$

Covariance: $\text{cov}(Y_i, Y_{i^*})=\sum_s P(Y_i = y_s; Y_{i^*} = y_{s^*})\left(y_s - E(Y_i)\right)\left(y_{s^*} - E(Y_{i^*})\right)$

Variance matrix: $\text{var}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \text{var}(Y_1) & \text{cov}(Y_1, Y_2) \\ \text{cov}(Y_1, Y_2) & \text{var}(Y_2) \end{pmatrix}$

Linear combinations of Random Variables:

$$T = aY_1 + bY_2 = \begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

$$E(T) = \begin{pmatrix} a & b \end{pmatrix} E\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix},$$

$$\text{var}(T) = \begin{pmatrix} a & b \end{pmatrix} \text{var}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

4. REPRESENTING A SAMPLE, RANDOM VARIABLES FOR SIMPLE RANDOM SAMPING (WITHOUT REPLACEMENT)

Permutation –

Position- (differs from label)

Indicator random variable for selection of unit in position U_{is}.

Representing the unit in position i: $S_i = \sum_{s=1}^{N} U_{is} s$

Representing random variable in position i: $Y_i = \sum_{s=1}^{N} U_{is} y_s$

Sample and Remainder

Define the sample as the first part of a permutation

Sample: $i = 1, \ldots, n$

Remainder: $i = n+1, \ldots, N$

Sampling fraction: $f = \frac{n}{N}$

Expected value and variance of U_{is}.

$$E\left(U_{is}\right) = \frac{1}{N}; \quad \text{var}\left(U_{is}\right) = \frac{1}{N}\left(1 - \frac{1}{N}\right);$$

Covariances of U_{is} with $U_{i^*s^*}$ for all i and s

When $s \neq s^*$, $E\left(U_{is}U_{is^*}\right) = 0$

When $i \neq i^*$, $E\left(U_{is}U_{i^*s}\right) = 0$

When $s \neq s^*$ and $i \neq i^*$, $E\left(U_{is}U_{i^*s^*}\right) = \frac{1}{N\left(N-1\right)}$

When $s \neq s^*$, $\text{cov}\left(U_{is}, U_{is^*}\right) = -\frac{1}{N^2}$

When $i \neq i^*$, $\text{cov}\left(U_{is}, U_{i^*s}\right) = -\frac{1}{N^2}$

When $s \neq s^*$ and $i \neq i^*$, $\text{cov}\left(U_{is}, U_{is^*}\right) = \frac{1}{\left(N-1\right)N^2}$

5. USING VECTORS AND MATRICES TO SIMPLIFY REPRESENTATION OF RANDOM VARIABLES FOR A RANDOM PERMUTATION

Basic notation- bold, column vectors:

$$\mathbf{I}_a = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}; \ \mathbf{1}_a = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix};$$

$$\mathbf{J}_a = \mathbf{1}_a \mathbf{1}_a' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix};$$

$$\mathbf{P}_a = \mathbf{I}_a - \frac{1}{a}\mathbf{J}_a,$$

Define:

Indicator Random Variables: $\mathbf{U}_i = \left(U_{i1} \quad U_{i2} \quad \cdots \quad U_{iN}\right)'$, Population Parameters: $\mathbf{y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_N \end{pmatrix}'$, and

$\mathbf{U}' = \left((\mathbf{U}_i) \right) = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_N \end{pmatrix}$.

Properties:

$$E(\mathbf{U}_i) = \frac{1}{N}\mathbf{1}_N;$$

$$\text{var}(\mathbf{U}_i) = \frac{1}{N}\mathbf{P}_N \text{ or } \text{var}(\mathbf{U}_i) = \frac{1}{N-1}\left(\mathbf{P}_N - \frac{1}{N}\mathbf{P}_N\right).$$

$$\text{cov}(\mathbf{U}_i, \mathbf{U}_{i*}) = \frac{1}{N-1}\left(-\frac{1}{N}\mathbf{P}_N\right) \text{ when } i \neq i*.$$

We use these results to develop expressions for the expected value and variance of random variables such as $\mathbf{Y} = \mathbf{U}\mathbf{y}$.

6. EXPECTED VALUE OF $\mathbf{U}$ AND VARIANCE OF $vec(\mathbf{U})$

Additional matrix algebra notation:

Let $\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_N \end{pmatrix}$ where $\mathbf{A}_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{iN} \end{pmatrix}'$

$$vec(\mathbf{A}) = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_N \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1' & \mathbf{A}_2' & \cdots & \mathbf{A}_N' \end{pmatrix}'$$

$$vec(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A}) vec(\mathbf{B})$$

Kronecker product $\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & a_{13}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & a_{23}\mathbf{B} \\ a_{31}\mathbf{B} & a_{31}\mathbf{B} & a_{31}\mathbf{B} \end{pmatrix}$, and

Direct sums $\bigoplus_{i=1}^{N} \mathbf{A}_i = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_N \end{pmatrix}$.

Re-express $\mathbf{Y}$

$\mathbf{Y} = vec(\mathbf{Y}')$ or

$\mathbf{Y} = vec(\mathbf{y}'\mathbf{U}')$ or

$\mathbf{Y} = (\mathbf{I}_N \otimes \mathbf{y}') vec(\mathbf{U}')$

Express $E\left[vec(\mathbf{U}')\right]$ and $\text{var}\left[vec(\mathbf{U}')\right]$

$E\left[vec(\mathbf{U}')\right] = E\left[vec\begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_N \end{pmatrix}\right]$. Then

$$E\left[vec\left(\mathbf{U}'\right)\right]=\frac{1}{N}\mathbf{1}_N\otimes\mathbf{1}_N \text{ , and}$$

$$\text{var}\left[vec\left(\mathbf{U}'\right)\right]=\frac{1}{N-1}\mathbf{P}_N\otimes\mathbf{P}_N \text{ since}$$

$$\text{var}\left[vec\left(\mathbf{U}'\right)\right]=\text{var}\begin{pmatrix}\mathbf{U}_1\\ \mathbf{U}_2\\ \vdots\\ \mathbf{U}_N\end{pmatrix}$$

$$=\frac{1}{N-1}\left[\begin{pmatrix}\mathbf{P}_N & \mathbf{0}_N & \cdots & \mathbf{0}_N\\ \mathbf{0}_N & \mathbf{P}_N & \cdots & \mathbf{0}_N\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0}_N & \mathbf{0}_N & \cdots & \mathbf{P}_N\end{pmatrix}-\frac{1}{N}\begin{pmatrix}\mathbf{P}_N & \mathbf{P}_N & \cdots & \mathbf{P}_N\\ \mathbf{P}_N & \mathbf{P}_N & \cdots & \mathbf{P}_N\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{P}_N & \mathbf{P}_N & \cdots & \mathbf{P}_N\end{pmatrix}\right].$$

$$=\frac{1}{N-1}\left[\mathbf{I}_N-\frac{1}{N}\mathbf{J}_N\right]\otimes\mathbf{P}_N$$

7. DEFINE THE EXPANDED MODEL i.e. $\overleftrightarrow{Y}_{is}=U_{is}y_s$.
(Note: Notation differs from that of Stanek, Singer, and Lencina (2004))

Define $\overleftrightarrow{\mathbf{Y}}_i=\left(\overset{N}{\underset{s=1}{\oplus}}y_s\right)\mathbf{U}_i=\left(U_{i1}y_1 \quad U_{i2}y_2 \quad \cdots \quad U_{iN}y_N\right)'$

Define the expanded random variables as

$$\overleftrightarrow{\mathbf{Y}}=vec\begin{pmatrix}U_{11}y_1 & U_{21}y_1 & \cdots & U_{N1}y_1\\ U_{12}y_2 & U_{22}y_2 & \cdots & U_{N2}y_2\\ \vdots & \vdots & \ddots & \vdots\\ U_{1N}y_N & U_{2N}y_N & \cdots & U_{NN}y_N\end{pmatrix} \text{ or}$$

$$\overleftrightarrow{\mathbf{Y}}=vec\left(\overleftrightarrow{\mathbf{Y}}_1 \quad \overleftrightarrow{\mathbf{Y}}_2 \quad \cdots \quad \overleftrightarrow{\mathbf{Y}}_N\right)' \text{ or}$$

$$\overleftrightarrow{\mathbf{Y}}=\left[\mathbf{I}_N\otimes\left(\overset{N}{\underset{s=1}{\oplus}}y_s\right)\right]vec\left(\mathbf{U}'\right).$$

$$E\left(\overleftrightarrow{\mathbf{Y}}\right)=\left[\frac{1}{N}\left(\mathbf{1}_N\otimes\mathbf{I}_N\right)\right]\mathbf{y}.$$

$$\text{var}\left(\overleftrightarrow{\mathbf{Y}}\right)=\mathbf{P}_N\otimes\mathbf{\Delta} \text{ where } \mathbf{\Delta}=\frac{1}{N-1}\left(\overset{N}{\underset{s=1}{\oplus}}y_s\right)\mathbf{P}_N\left(\overset{N}{\underset{s=1}{\oplus}}y_s\right).$$

8. COLLAPSING THE EXPANDED RANDOM VARIABLES

Linear combinations of the expanded random variables correspond to constants and random variables that are commonly used. We represent collapsed random variables via expressions of the form: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\left(\mathbf{A}\otimes\mathbf{B}\right)\overleftrightarrow{\mathbf{Y}}$. To understand the collapsed set of random variables, it is often helpful to visualize the equivalent expression:

$$\mathbf{L}\overleftrightarrow{\mathbf{Y}}=vec\left(\mathbf{B}\begin{pmatrix} U_{11}y_1 & U_{21}y_1 & \cdots & U_{N1}y_1 \\ U_{12}y_2 & U_{22}y_2 & \cdots & U_{N2}y_2 \\ \vdots & \vdots & \ddots & \vdots \\ U_{1N}y_N & U_{2N}y_N & \cdots & U_{NN}y_N \end{pmatrix}\mathbf{A}'\right).$$

a. Collapsing to $\mathbf{y}$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\mathbf{y}$ where $\mathbf{L}=\mathbf{1}_N'\otimes\mathbf{I}_N$ (population frame)

b. Collapsing to $\mathbf{Y}$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\mathbf{Y}$ where $\mathbf{L}=\mathbf{I}_N\otimes\mathbf{1}_N'$ (usual random variables)

c. Collapsing to sample and remainder totals $\begin{pmatrix} Y_I \\ Y_{II} \end{pmatrix}$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\begin{pmatrix} Y_I \\ Y_{II} \end{pmatrix}$ where $\mathbf{L}=\begin{pmatrix} \mathbf{1}_n' & \mathbf{0}_{N-n}' \\ \mathbf{0}_n' & \mathbf{1}_{N-n}' \end{pmatrix}\otimes\mathbf{1}_N'$

where $Y_I=\sum_{i=1}^{n}Y_i$ and $Y_{II}=\sum_{i=n+1}^{N}Y_i$.

d. Collapsing to sample and remainder means $\begin{pmatrix} \bar{Y}_I \\ \bar{Y}_{II} \end{pmatrix}$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\begin{pmatrix} \bar{Y}_I \\ \bar{Y}_{II} \end{pmatrix}$ where

$$\mathbf{L}=\frac{1}{N}\begin{pmatrix} \frac{1}{f}\mathbf{1}_n' & \mathbf{0}_{N-n}' \\ \mathbf{0}_n' & \frac{1}{1-f}\mathbf{1}_{N-n}' \end{pmatrix}\otimes\mathbf{1}_N' \text{ and } \bar{Y}_I=\frac{1}{n}\sum_{i=1}^{n}Y_i \text{ and } \bar{Y}_{II}=\frac{1}{N-n}\sum_{i=n+1}^{N}Y_i .$$

e. Collapsing to weighted population frame: $\mathbf{y}_w$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\mathbf{y}_w$ where $\mathbf{L}=\mathbf{1}_N'\otimes\left(\bigoplus_{s=1}^{N}w_s\right)$ where

$\sum_{s=1}^{N}w_s=1$, $\mathbf{y}_w=\left(\left(y_{ws}\right)\right)$ and $y_{ws}=w_s y_s$.

f. Collapsing to weighted population random variables: $\mathbf{Y}_w$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\mathbf{Y}_w$ where $\mathbf{L}=\mathbf{I}_N\otimes\mathbf{w}'$

where $\mathbf{w}=\left(\left(w_s\right)\right)=\begin{pmatrix} w_1 & w_2 & \cdots & w_N \end{pmatrix}'$, $\sum_{s=1}^{N}w_s=1$, $\mathbf{Y}_w=\left(\left(Y_{wi}\right)\right)$ and $Y_{wi}=\sum_{s=1}^{N}U_{is}y_{ws}$.

g. Collapsing to weighted sample and remainder totals $\begin{pmatrix} Y_{wI} \\ Y_{wII} \end{pmatrix}$: $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\begin{pmatrix} Y_{wI} \\ Y_{wII} \end{pmatrix}$ where

$$\mathbf{L}=\begin{pmatrix}\mathbf{1}_n' & \mathbf{0}_{N-n}' \\ \mathbf{0}_n' & \mathbf{1}_{N-n}'\end{pmatrix}\otimes\mathbf{w}' \text{ where } \mathbf{w}=\left(\left(w_s\right)\right)=\begin{pmatrix}w_1 & w_2 & \cdots & w_N\end{pmatrix}',\ \sum_{s=1}^{N}w_s=1,\ \mathbf{Y}_w=\left(\left(Y_{wi}\right)\right),$$

$$Y_{wi}=\sum_{s=1}^{N}U_{is}w_s y_s,\ Y_{wI}=\sum_{i=1}^{n}Y_{wi} \text{ and } Y_{wII}=\sum_{i=n+1}^{N}Y_{wi}.$$

Sufficiency: It is valuable to consider whether or not a collapsed set of random variables is sufficient for a linear combination of $\overleftrightarrow{\mathbf{Y}}$. If it is sufficient, we can develop inference based on $\mathbf{L}\overleftrightarrow{\mathbf{Y}}$ that will be optimal. Otherwise, the inference based solely on $\mathbf{L}\overleftrightarrow{\mathbf{Y}}$ will not be optimal.

Sufficiency has been discussed only with respect to linear combinations of $\mathbf{Y}$ (Stanek and Singer, 2004, referring to Rao and Bellhouse, 1978, Theorem 1.1). Briefly, express

$$\mathbf{I}_N\otimes\mathbf{I}_N=\mathbf{I}_N\otimes\mathbf{P}_N+\frac{1}{N}\left(\mathbf{I}_N\otimes\mathbf{1}_N\right)\left(\mathbf{I}_N\otimes\mathbf{1}_N'\right).$$

Then $\overleftrightarrow{\mathbf{Y}}=\left(\mathbf{I}_N\otimes\mathbf{P}_N\right)\overleftrightarrow{\mathbf{Y}}+\frac{1}{N}\left(\mathbf{I}_N\otimes\mathbf{1}_N\right)\mathbf{Y}$ where $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\mathbf{Y}$ and $\mathbf{L}=\mathbf{I}_N\otimes\mathbf{1}_N'$. We have claimed (see Stanek, Singer and Lencina (2004)) that $\mathbf{Y}$ is sufficient since $\left(\mathbf{I}_N\otimes\mathbf{P}_N\right)$ and $\frac{1}{N}\left(\mathbf{I}_N\otimes\mathbf{1}_N\right)$ are orthogonal, and the expected value of $\left(\mathbf{I}_N\otimes\mathbf{P}_N\right)\overleftrightarrow{\mathbf{Y}}$ is zero (see Rao and Bellhouse, 1978, Theorem 1.1). In other documents in this report, we investigate this theorem in the context of measurement error, and in two stage sampling.

9. PARTITIONING INTO THE SAMPLE AND REMAINDER

Collapsed random variables that are a linear combination of $\mathbf{Y}$ or $\mathbf{Y}_w$ can be partitioned into a sample and a remainder (collapsing b-e in Section 8) by partitioning $\mathbf{L}=\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}$. In general, we represent the result as $\mathbf{L}\overleftrightarrow{\mathbf{Y}}=\begin{pmatrix}\mathbf{L}_I\overleftrightarrow{\mathbf{Y}}\\ \mathbf{L}_{II}\overleftrightarrow{\mathbf{Y}}\end{pmatrix}=\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}$. Using properties of $\overleftrightarrow{\mathbf{Y}}$, we evaluate $E\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{X}_I\\ \mathbf{X}_{II}\end{pmatrix}\boldsymbol{\gamma}$ (where $\boldsymbol{\gamma}$ represents some parameters) and $\text{var}\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II}\\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}$. We illustrate this for collapsing b-e in Section 8.

b. Usual random variables $\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}$: $\mathbf{L}=\mathbf{I}_N\otimes\mathbf{1}_N'$ and $\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}\mathbf{I}_n & \underset{n\times N-n}{\mathbf{0}}\end{pmatrix}\otimes\mathbf{1}_N'\\ \hline \begin{pmatrix}\underset{(N-n)xn}{\mathbf{0}} & \mathbf{I}_{N-n}\end{pmatrix}\otimes\mathbf{1}_N'\end{pmatrix}$ where $\mathbf{Y}_I=\left(\left(Y_i\right)\right)=\begin{pmatrix}Y_1 & Y_2 & \cdots & Y_n\end{pmatrix}'$ and $\mathbf{Y}_{II}=\left(\left(Y_i\right)\right)=\begin{pmatrix}Y_{n+1} & Y_{n+2} & \cdots & Y_N\end{pmatrix}'$. Then,

$$E\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{1}_n\\ \mathbf{1}_{N-n}\end{pmatrix}\mu \text{ and } \operatorname{var}\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}=\left(\begin{array}{c|c}\mathbf{I}_n-\frac{1}{N}\mathbf{J}_n & -\frac{1}{N}\mathbf{J}_{n\times(N-n)}\\ \hline -\frac{1}{N}\mathbf{J}_{(N-n)\times n} & \mathbf{I}_{N-n}-\frac{1}{N}\mathbf{J}_{N-n}\end{array}\right)\sigma^2 \text{ where } \mathbf{1}_N'\boldsymbol{\Delta}\mathbf{1}_N=\sigma^2.$$

c. Collapsing to sample and remainder totals $\begin{pmatrix}Y_I\\ Y_{II}\end{pmatrix}$: $\mathbf{L}=\begin{pmatrix}\mathbf{1}_n' & \mathbf{0}_{N-n}'\\ \mathbf{0}_n' & \mathbf{1}_{N-n}'\end{pmatrix}\otimes\mathbf{1}_N'$ where

$\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}=\left(\begin{array}{c}(\mathbf{1}_n' \ \ \mathbf{0}_{N-n}')\otimes\mathbf{1}_N'\\ \hline (\mathbf{0}_n' \ \ \mathbf{1}_{N-n}')\otimes\mathbf{1}_N'\end{array}\right)$. Then $E\begin{pmatrix}Y_I\\ Y_{II}\end{pmatrix}=N\begin{pmatrix}f\\ 1-f\end{pmatrix}\mu$ and

$\operatorname{var}\begin{pmatrix}Y_I\\ Y_{II}\end{pmatrix}=Nf(1-f)\left(\begin{array}{c|c}1 & -1\\ \hline -1 & 1\end{array}\right)\sigma^2$.

d. Collapsing to sample and remainder means $\begin{pmatrix}\bar{Y}_I\\ \bar{Y}_{II}\end{pmatrix}$: $\mathbf{L}=\frac{1}{N}\begin{pmatrix}\frac{1}{f}\mathbf{1}_n' & \mathbf{0}_{N-n}'\\ \mathbf{0}_n' & \frac{1}{1-f}\mathbf{1}_{N-n}'\end{pmatrix}\otimes\mathbf{1}_N'$ where

$\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}=\left(\begin{array}{c}\frac{1}{N}\left(\frac{1}{f}\mathbf{1}_n' \ \ \mathbf{0}_{N-n}'\right)\otimes\mathbf{1}_N'\\ \hline \frac{1}{N}\left(\mathbf{0}_n' \ \ \frac{1}{1-f}\mathbf{1}_{N-n}'\right)\otimes\mathbf{1}_N'\end{array}\right)$. Then $E\begin{pmatrix}\bar{Y}_I\\ \bar{Y}_{II}\end{pmatrix}=\begin{pmatrix}1\\ 1\end{pmatrix}\mu$ and

$\operatorname{var}\begin{pmatrix}\bar{Y}_I\\ \bar{Y}_{II}\end{pmatrix}=\frac{1}{n(1-f)}\left(\begin{array}{c|c}(1-f)^2 & -f(1-f)\\ \hline -f(1-f) & f^2\end{array}\right)\sigma^2$.

e. Collapsing to the population frame $\begin{pmatrix}\mathbf{y}_{wI}\\ \mathbf{y}_{wII}\end{pmatrix}$: $\mathbf{L}=\mathbf{1}_N'\otimes\left(\bigoplus_{s=1}^{N}w_s\right)$ where

$\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}=\left(\begin{array}{c}\mathbf{1}_N'\otimes\left(\bigoplus_{s=1}^{n}w_s \ \ \mathbf{0}_{n\times N-n}\right)\\ \hline \mathbf{1}_N'\otimes\left(\mathbf{0}_{(N-n)xn} \ \ \bigoplus_{s=n+1}^{N}w_s\right)\end{array}\right)$ and where $\mathbf{y}_{wI}=((y_{ws}))=(y_{w1} \ \ y_{w2} \ \ \cdots \ \ y_{wn})'$ and

$\mathbf{y}_{wII}=((y_{ws}))=(y_{w(n+1)} \ \ y_{w(n+2)} \ \ \cdots \ \ y_{wN})'$. Then $E\left[\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}\overleftrightarrow{\mathbf{Y}}\right]=\left(\begin{array}{c}(\mathbf{I}_n \ \ \mathbf{0}_{n\times N-n})\\ \hline (\mathbf{0}_{(N-n)xn} \ \ \mathbf{I}_{N-n})\end{array}\right)\begin{pmatrix}\mathbf{y}_{wI}\\ \mathbf{y}_{wII}\end{pmatrix}$, and

$\operatorname{var}\left[\begin{pmatrix}\mathbf{L}_I\\ \mathbf{L}_{II}\end{pmatrix}\overleftrightarrow{\mathbf{Y}}\right]=\mathbf{0}$.

f. Collapsing to weighted population random variables: $\begin{pmatrix} \mathbf{Y}_{wI} \\ \mathbf{Y}_{wII} \end{pmatrix}$ where

$\mathbf{Y}_{wI} = \left(\left(Y_{wi}\right)\right) = \begin{pmatrix} Y_{w1} & Y_{w2} & \cdots & Y_{wn} \end{pmatrix}'$ and $\mathbf{Y}_{wII} = \left(\left(Y_{wi}\right)\right) = \begin{pmatrix} Y_{w(n+1)} & Y_{w(n+2)} & \cdots & Y_{wN} \end{pmatrix}'$:

$\mathbf{L} = \mathbf{I}_N \otimes \mathbf{w}'$ where $\begin{pmatrix} \mathbf{L}_I \\ \mathbf{L}_{II} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \mathbf{I}_n & \underset{n\times N-n}{\mathbf{0}} \end{pmatrix} \otimes \mathbf{w}' \\ \hline \begin{pmatrix} \underset{(N-n)xn}{\mathbf{0}} & \mathbf{I}_{N-n} \end{pmatrix} \otimes \mathbf{w}' \end{pmatrix}$. $E\begin{pmatrix} \mathbf{Y}_{wI} \\ \mathbf{Y}_{wII} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n \\ \mathbf{1}_{N-n} \end{pmatrix} \mu_w$ where we define

$\mu_w = \frac{1}{N}\sum_{s=1}^{N} w_s y_s$. Also, $\text{var}\begin{pmatrix} \mathbf{Y}_{wI} \\ \mathbf{Y}_{wII} \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I}_n - \frac{1}{N}\mathbf{J}_n & -\frac{1}{N}\underset{n\times(N-n)}{\mathbf{J}} \\ \hline -\frac{1}{N}\underset{(N-n)\times n}{\mathbf{J}} & \mathbf{I}_{N-n} - \frac{1}{N}\mathbf{J}_{N-n} \end{array}\right)\sigma_w^2$ where $\sigma_w^2 = \mathbf{w}'\boldsymbol{\Delta}\mathbf{w}$ or

$\sigma_w^2 = \frac{1}{N-1}\sum_{s=1}^{N}\left(w_s y_s - \mu_w\right)^2$.

g. Collapsing to weighted sample and remainder totals $\begin{pmatrix} Y_{wI} \\ Y_{wII} \end{pmatrix}$: $\mathbf{L} = \begin{pmatrix} \mathbf{1}_n' & \mathbf{0}_{N-n}' \\ \mathbf{0}_n' & \mathbf{1}_{N-n}' \end{pmatrix} \otimes \mathbf{w}'$ where

$\begin{pmatrix} \mathbf{L}_I \\ \mathbf{L}_{II} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \mathbf{1}_n' & \mathbf{0}_{N-n}' \end{pmatrix} \otimes \mathbf{w}' \\ \hline \begin{pmatrix} \mathbf{0}_n' & \mathbf{1}_{N-n}' \end{pmatrix} \otimes \mathbf{w}' \end{pmatrix}$ where $\mathbf{w} = \left(\left(w_s\right)\right) = \begin{pmatrix} w_1 & w_2 & \cdots & w_N \end{pmatrix}'$ and $\sum_{s=1}^{N} w_s = 1$. Then

$E\begin{pmatrix} Y_{wI} \\ Y_{wII} \end{pmatrix} = N\begin{pmatrix} f \\ \hline 1-f \end{pmatrix}\mu_w$ and $\text{var}\begin{pmatrix} Y_{wI} \\ Y_{wII} \end{pmatrix} = Nf\left(1-f\right)\left(\begin{array}{c|c} 1 & -1 \\ \hline -1 & 1 \end{array}\right)\sigma_w^2$.

10. TARGET PARAMETERS, RANDOM VARIABLES, AND TERMINOLOGY

We assume that there is an interest in a ‘target’ that can be defined as a linear combination of the collapsed random variables. We use the sample data to estimate/predict the target.

If the target is a fixed constant, then we use the sample data to “estimate” the target.
If the target is a random variable, then we use the sample data to “predict” the target.

We only consider inference for a single ‘target’ (not joint targets). We assume
$E\begin{pmatrix} \mathbf{Y}_I \\ \mathbf{Y}_{II} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_I \\ \mathbf{X}_{II} \end{pmatrix}\boldsymbol{\alpha}$.

We represent collapsed random variables by $\mathbf{L}\vec{\mathbf{Y}} = \begin{pmatrix} \mathbf{Y}_I \\ \mathbf{Y}_{II} \end{pmatrix}$. We define a ‘target’ as $P = \mathbf{g}'\begin{pmatrix} \mathbf{Y}_I \\ \mathbf{Y}_{II} \end{pmatrix}$ and partition $\mathbf{g}$ representing $\mathbf{g}' = \begin{pmatrix} \mathbf{g}_I' & \mathbf{g}_{II}' \end{pmatrix}$ so that $P = \mathbf{g}_I'\mathbf{Y}_I + \mathbf{g}_{II}'\mathbf{Y}_{II}$. Since $\mathbf{Y}_I$ will be realized after selecting the sample and observing the results, the basic inference problem is prediction of

$\mathbf{g}_{II}'\mathbf{Y}_{II}$. This is true when P is a fixed constant, and when P is a random variable. For this reason, we use the term 'predictor' when discussing inference.

Example: Suppose we collapse random variables to sample and remainder means $\begin{pmatrix} \bar{Y}_I \\ \bar{Y}_{II} \end{pmatrix}$, and are interested in the target $P = \mu$, a fixed constant. Since $P = f\bar{Y}_I + (1-f)\bar{Y}_{II}$, $\mathbf{g}' = \begin{pmatrix} f & (1-f) \end{pmatrix}$ where $g_I = f$ and $g_{II} = 1-f$.

11. DEVELOPING THE BEST LINEAR UNBIASED PREDICTOR (BLUP)

We define the BLUP of P as $\hat{P}$, where $\hat{P}$ satisfies the following criteria (see Royall, 1976):

Linear in the sample: $\hat{P} = (\mathbf{g}_I' + \mathbf{a}')\mathbf{Y}_I$

Unbiased: $E(\hat{P} - P) = 0$

Minimum MSE: $\text{var}(\hat{P} - P)$ is minimized.

The Unbiased Constraint:

We can expand the unbiased constraint: $E(\hat{P} - P) = \begin{pmatrix} \mathbf{a}' & -\mathbf{g}_{II}' \end{pmatrix}\begin{pmatrix} \mathbf{X}_I \\ \mathbf{X}_{II} \end{pmatrix}\boldsymbol{\alpha}$, such that $E(\hat{P} - P) = (\mathbf{a}'\mathbf{X}_I - \mathbf{g}_{II}'\mathbf{X}_{II})\boldsymbol{\alpha}$. In order for this expression to equal zero for any value of $\boldsymbol{\alpha}$, the unbiased constraint will be always be satisfied when $\mathbf{a}'\mathbf{X}_I - \mathbf{g}_{II}'\mathbf{X}_{II} = \mathbf{0}$. This is introduced as a constraint using Lagrangian multipliers when minimizing the MSE.

Finding the Minimum MSE:

Notice that $\text{var}(\hat{P} - P) = \text{var}\left[\begin{pmatrix} \mathbf{a}' & -\mathbf{g}_{II}' \end{pmatrix}\begin{pmatrix} \mathbf{Y}_I \\ \mathbf{Y}_{II} \end{pmatrix}\right] = \begin{pmatrix} \mathbf{a}' & -\mathbf{g}_{II}' \end{pmatrix}\begin{pmatrix} \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II} \end{pmatrix}\begin{pmatrix} \mathbf{a} \\ -\mathbf{g}_{II} \end{pmatrix}$.

Expanding this expression, $\text{var}(\hat{P} - P) = \mathbf{a}'\mathbf{V}_I\mathbf{a} - 2\mathbf{g}_{II}'\mathbf{V}_{II,I}\mathbf{a} + \mathbf{g}_{II}'\mathbf{V}_{II}\mathbf{g}_{II}$. Including the constraint via a Lagrangian multiplier, we seek to find the value of $\mathbf{a}$ that will minimize

$$f(\mathbf{a}, \boldsymbol{\lambda}) = \mathbf{a}'\mathbf{V}_I\mathbf{a} - 2\mathbf{g}_{II}'\mathbf{V}_{II,I}\mathbf{a} + \mathbf{g}_{II}'\mathbf{V}_{II}\mathbf{g}_{II} + 2(\mathbf{a}'\mathbf{X}_I - \mathbf{g}_{II}'\mathbf{X}_{II})\boldsymbol{\lambda}.$$

Differentiating with respect to $\mathbf{a}$ and $\boldsymbol{\lambda}$,

$$\frac{\partial f(\mathbf{a}, \boldsymbol{\lambda})}{\partial \mathbf{a}} = 2\mathbf{V}_I\mathbf{a} - 2\mathbf{V}_{I,II}\mathbf{g}_{II} + 2\mathbf{X}_I\boldsymbol{\lambda} \text{ and}$$

$$\frac{\partial f(\mathbf{a}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = 2(\mathbf{X}_I'\mathbf{a} - \mathbf{X}_{II}'\mathbf{g}_{II}).$$

To find the value of $\mathbf{a}$ that minimizes $f\left(\mathbf{a},\boldsymbol{\lambda}\right)$, we set these derivatives to zero simultaneously, and solve for $\mathbf{a}$. The estimating equations are given by

$$\frac{1}{2}\begin{pmatrix} \dfrac{\partial f\left(\hat{\mathbf{a}},\hat{\boldsymbol{\lambda}}\right)}{\partial \mathbf{a}} \\ \dfrac{\partial f\left(\hat{\mathbf{a}},\hat{\boldsymbol{\lambda}}\right)}{\partial \boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \mathbf{V}_I & \mathbf{X}_I \\ \mathbf{X}_I' & \mathbf{0} \end{pmatrix}\begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} - \begin{pmatrix} \mathbf{V}_{I,II}\mathbf{g}_{II} \\ \mathbf{X}_{II}'\mathbf{g}_{II} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_n \\ \mathbf{0} \end{pmatrix}$$

or equivalently, by $\begin{pmatrix} \mathbf{V}_I & \mathbf{X}_I \\ \mathbf{X}_I' & \mathbf{0} \end{pmatrix}\begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \mathbf{V}_{I,II}\mathbf{g}_{II} \\ \mathbf{X}_{II}'\mathbf{g}_{II} \end{pmatrix}$.

To solve this equation for $\hat{\mathbf{a}}$, we make use of results on the inverse of a partitioned matrix $\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$, such that $\mathbf{X}^{-1} = \left(\begin{array}{c|c} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{Q}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{Q}^{-1} \\ \hline -\mathbf{Q}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{Q}^{-1} \end{array}\right)$ where $\mathbf{Q} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$. As a result,

$$\begin{pmatrix} \mathbf{V}_I & \mathbf{X}_I \\ \mathbf{X}_I{}' & 0 \end{pmatrix}^{-1} = \left(\begin{array}{c|c} \mathbf{V}_I^{-1} - \mathbf{V}_I^{-1}\mathbf{X}_I\left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I{}'\mathbf{V}_I^{-1} & \mathbf{V}_I^{-1}\mathbf{X}_I\left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1} \\ \hline \left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I{}'\mathbf{V}_I^{-1} & -\left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1} \end{array}\right),$$

$$\hat{\mathbf{a}} = \left[\mathbf{V}_I^{-1} - \mathbf{V}_I^{-1}\mathbf{X}_I\left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I{}'\mathbf{V}_I^{-1}\right]\mathbf{V}_{I,II}\mathbf{g}_{II} + \mathbf{V}_I^{-1}\mathbf{X}_I\left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_{II}{}'\mathbf{g}_{II}.$$

We can now express the best linear unbiased predictor. Recall that $P = \mathbf{g}_I{}'\mathbf{Y}_I + \mathbf{g}_{II}{}'\mathbf{Y}_{II}$, and we predict $\mathbf{g}_{II}{}'\mathbf{Y}_{II}$ by $\hat{\mathbf{a}}'\mathbf{Y}_I$. Let us define $\hat{\alpha} = \left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{Y}_I$. Then

$$\begin{aligned} \hat{\mathbf{a}}'\mathbf{Y}_I &= \mathbf{g}_{II}{}'\mathbf{V}_{I,II}{}'\mathbf{V}_I^{-1}\mathbf{Y}_I - \mathbf{g}_{II}{}'\mathbf{V}_{I,II}{}'\mathbf{V}_I^{-1}\mathbf{X}_I\hat{\alpha} + \mathbf{g}_{II}{}'\mathbf{X}_{II}\hat{\alpha} \\ &= \mathbf{g}_{II}{}'\left[\mathbf{X}_{II}\hat{\alpha} + \mathbf{V}_{I,II}{}'\mathbf{V}_I^{-1}\left(\mathbf{Y}_I - \mathbf{X}_I\hat{\alpha}\right)\right] \end{aligned}.$$

As a result, the best linear unbiased predictor is given by

$$\hat{P} = \mathbf{g}_I{}'\mathbf{Y}_I + \mathbf{g}_{II}{}'\left[\mathbf{X}_{II}\hat{\alpha} + \mathbf{V}_{II,I}\mathbf{V}_I^{-1}\left(\mathbf{Y}_I - \mathbf{X}_I\hat{\alpha}\right)\right].$$

<u>Example 1.</u>

Target: $\mu = \frac{1}{N}\sum_{i=1}^{N} Y_i$

Collapsing: $\mathbf{L}\vec{\mathbf{Y}} = \mathbf{Y}$ (usual random variables)

$$E\begin{pmatrix} \mathbf{Y}_I \\ \mathbf{Y}_{II} \end{pmatrix} = \left(\begin{array}{c} \mathbf{1}_n \\ \hline \mathbf{1}_{N-n} \end{array}\right)\mu \text{ and } \begin{pmatrix} \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II} \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{I}_n - \frac{1}{N}\mathbf{J}_n & -\frac{1}{N}\mathbf{J}_{n\times(N-n)} \\ \hline -\frac{1}{N}\mathbf{J}_{(N-n)\times n} & \mathbf{I}_{N-n} - \frac{1}{N}\mathbf{J}_{N-n} \end{array}\right)\sigma^2.$$

Now $\begin{pmatrix}\mathbf{g}_I' & \mathbf{g}_{II}'\end{pmatrix}=\begin{pmatrix}\frac{1}{N}\mathbf{1}_n' & \frac{1}{N}\mathbf{1}_{N-n}'\end{pmatrix}$ so that the unbiased constraint simplifies to $\mathbf{a}'\mathbf{1}_n=1-f$. To evaluate the predictor, we note that $\mathbf{V}_I^{-1}=\left(\mathbf{I}_n-\frac{1}{N}\mathbf{J}_n\right)^{-1}=\mathbf{I}_n+\frac{1}{N-n}\mathbf{J}_n$. Other terms simplify: $\hat{\alpha}=\left(\mathbf{X}_I'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{-1}\mathbf{Y}_I=\frac{1}{n}\sum_{i=1}^{n}Y_i=\overline{Y}_I$; $\mathbf{V}_{II,I}\mathbf{V}_I^{-1}=-\left(\frac{1}{N-n}\mathbf{1}_{N-n}\right)\mathbf{1}_n'$ so that

$$
\begin{aligned}
\hat{P}&=\mathbf{g}_I'\mathbf{Y}_I+\mathbf{g}_{II}'\left[\mathbf{X}_{II}\hat{\alpha}+\mathbf{V}_{II,I}\mathbf{V}_I^{-1}\left(\mathbf{Y}_I-\mathbf{X}_I\hat{\alpha}\right)\right]\\
&=\frac{1}{N}\mathbf{1}_n'\mathbf{Y}_I+\frac{1}{N}\mathbf{1}_{N-n}'\left[\mathbf{1}_{N-n}\overline{Y}_I+-\left(\frac{1}{N-n}\mathbf{1}_{N-n}\right)\mathbf{1}_n'\left(\mathbf{Y}_I-\mathbf{1}_n\overline{Y}_I\right)\right]\\
&=\frac{n}{N}\overline{Y}_I+\frac{N-n}{N}\left[\overline{Y}_I-\left(\frac{1}{N-n}\right)\mathbf{1}_n'\left(\mathbf{Y}_I-\mathbf{1}_n\overline{Y}_I\right)\right]\\
&=f\overline{Y}_I+(1-f)\overline{Y}_I\\
&=\overline{Y}_I
\end{aligned}
$$

Example 2.

Target: $\mu=f\overline{Y}_I+(1-f)\overline{Y}_{II}$

Collapse random variables to sample and remainder means $\begin{pmatrix}\overline{Y}_I\\ \overline{Y}_{II}\end{pmatrix}$

$$E\begin{pmatrix}\overline{Y}_I\\ \overline{Y}_{II}\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}\mu \text{ and } \operatorname{var}\begin{pmatrix}\overline{Y}_I\\ \overline{Y}_{II}\end{pmatrix}=\frac{1}{n(1-f)}\left(\begin{array}{c|c}(1-f)^2 & -f(1-f)\\ \hline -f(1-f) & f^2\end{array}\right)\sigma^2.$$

Now $\begin{pmatrix}\mathbf{g}_I' & \mathbf{g}_{II}'\end{pmatrix}=\begin{pmatrix}f & (1-f)\end{pmatrix}$ so that the unbiased constraint simplifies to $a=1-f$. There is no other choice for a that will be unbiased, so we set $\hat{a}=1-f$. As a result, the BLUP is given by

$$
\begin{aligned}
\hat{P}&=\left(\mathbf{g}_I'+\hat{\mathbf{a}}'\right)\mathbf{Y}_I\\
&=\left(f+(1-f)\right)\overline{Y}_I.\\
&=\overline{Y}_I
\end{aligned}
$$

Example 3.

Target: Y_i

Collapsing: $\mathbf{L}\vec{\mathbf{Y}}=\mathbf{Y}$ (usual random variables)

$$E\begin{pmatrix}\mathbf{Y}_I\\ \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{1}_n\\ \hline \mathbf{1}_{N-n}\end{pmatrix}\mu \text{ and } \begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II}\\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}=\left(\begin{array}{c|c}\mathbf{I}_n-\frac{1}{N}\mathbf{J}_n & -\frac{1}{N}\mathbf{J}_{n\times(N-n)}\\ \hline -\frac{1}{N}\mathbf{J}_{(N-n)\times n} & \mathbf{I}_{N-n}-\frac{1}{N}\mathbf{J}_{N-n}\end{array}\right)\sigma^2.$$

Now $\begin{pmatrix}\mathbf{g}_I' & \mathbf{g}_{II}'\end{pmatrix}=\begin{pmatrix}\mathbf{e}_{iI}' & \mathbf{e}_{iII}'\end{pmatrix}$, where $\mathbf{e}_i$ is an $N\times 1$ vector with element i equal to one, and all other elements equal to zero, partitioned to conform. The unbiased constraint simplifies to

$\mathbf{a}'\mathbf{1}_n = \mathbf{e}'_{iII}\mathbf{1}_{N-n}$. To evaluate the predictor, we note that $\mathbf{V}_I^{-1} = \left(\mathbf{I}_n - \frac{1}{N}\mathbf{J}_n\right)^{-1} = \mathbf{I}_n + \frac{1}{N-n}\mathbf{J}_n$.

Other terms simplify: $\hat{\alpha} = \left(\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I{}'\mathbf{V}_I^{-1}\mathbf{Y}_I = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}_I$; $\mathbf{V}_{II,I}\mathbf{V}_I^{-1} = -\left(\frac{1}{N-n}\mathbf{1}_{N-n}\right)\mathbf{1}'_n$ so that

$$\begin{aligned}
\hat{P} &= \mathbf{g}_I{}'\mathbf{Y}_I + \mathbf{g}_{II}{}'\left[\mathbf{X}_{II}\hat{\alpha} + \mathbf{V}_{II,I}\mathbf{V}_I^{-1}\left(\mathbf{Y}_I - \mathbf{X}_I\hat{\alpha}\right)\right] \\
&= \mathbf{e}'_{iI}\mathbf{Y}_I + \mathbf{e}'_{iII}\left[\mathbf{1}_{N-n}\overline{Y}_I + -\left(\frac{1}{N-n}\mathbf{1}_{N-n}\right)\mathbf{1}'_n\left(\mathbf{Y}_I - \mathbf{1}_n\overline{Y}_I\right)\right]. \\
&= \mathbf{e}'_{iI}\mathbf{Y}_I + \mathbf{e}'_{iII}\left[\mathbf{1}_{N-n}\overline{Y}_I\right] \\
&= \mathbf{e}'_{iI}\mathbf{Y}_I + \left(\mathbf{e}'_{iII}\mathbf{1}_{N-n}\right)\overline{Y}_I
\end{aligned}$$

Thus, when $i \le n$, $\hat{P} = Y_i$. When $i > n$, $\hat{P} = \overline{Y}_I$.