STT 825 question 7 chapter 12.
There may be a bonus question basic to this material on Exam 2.

As assumed in the statement of problem 7, chapter 12, the secondary unit scores \( yy(j), 1 \leq j \leq M \) are sampled iid from some probability distribution having mean \( \beta \) and variance \( \gamma \). It follows that their joint probability distribution is invariant under \textit{uniform random permutation} of secondary units (i.e. the joint distribution of \( yy(j), j \leq M \) is \textit{exchangeable}).

We may just as well dispense with primary unit sampling and select primary unit one, taking as our estimate of population total \( \tau \) the quantity \( N y(1) \) where \( y(1) \) denotes the \textbf{TOTAL} score of all \( m \) iid secondary units included in primary unit 1 and there are \( N \) primary units.

**Conditional interpretation of model unbiased estimator.** Denote by \( \mathcal{L} \) the \textbf{unordered} numbers \( yy(j), j \leq M \). As mentioned above, conditional on \( \mathcal{L} \), we have an \textit{exchangeable} model for \( yy \) and so

\[
E(N y(1) \mid \mathcal{L}) = E(y(1) \mid \mathcal{L}) + \ldots + E(y(1) \mid \mathcal{L}) = E(y(1) + \ldots + y(N) \mid \mathcal{L}) = E(\tau \mid \mathcal{L}) = \tau.
\]

Therefore with iid sampling of \( yy \) scores estimator \( N y(1) \) is a (model) \textit{conditionally unbiased estimator} of the \textit{random population total} \( \tau \).

**Conditional mean of \( N y(1) \).**

\[
E N y(1) = E E(N y(1) \mid \mathcal{L}) = E \tau = M \beta.
\]

Therefore this estimator is (model) \textit{unconditionally unbiased for the random population total} \( \tau \).

Consult the definition of conditional and unconditional unbiased pg. 133.

**Conditional variance of \( N y(1) \).** Conditional on \( \mathcal{L} \) r.v. \( y(1) \) is the sum of \( yy \) scores attached to a SRS of \( m \) from \( M yy \) scores from \( \mathcal{L} \) so

\[
\text{Var}(N y(1) \mid \mathcal{L}) = N^2 m^2 \text{Var}(y(1)/m \mid \mathcal{L}) = M^2 (1-m/M) s^2_{allyy} / m = M (N-1) s^2_{allyy} \text{ (recall } M = Nm \text{).} 
\]
Conditionally (model) unbiased estimator of \( \text{Var} (N \, y(1) \mid L) \).

\[
\hat{\text{Var}}(N \, y(1) \mid L) := M \, (N-1) \, s_{yy(j), \, j \, \in \, PU(1)}^2
\]

Conditionally on \( L \) the sample variance \( s_{yy(j), \, j \, \in \, PU(1)}^2 \) is that of a SRS of \( m \) from \( L \) and so \( E(M \, (N-1) \, s_{yy(j), \, j \, \in \, PU(1)}^2 \mid L) = \text{Var}(N \, y(1) \mid L) \).

**Conditional (model) zCI for \( \tau \).** We simply condition on \( L \) and use the estimator \( N \, y(1) \) which is conditionally unbiased for the conditionally constant population total \( \tau \), and the conditionally unbiased estimator of its conditional variance derived above. Subject to the usual caveats,

\[
N \, y(1) \pm z \, \sqrt{M(N - 1) \, s_{yy(j), \, j \, \in \, PU(1)}^2}
\]

**Unconditional (model) variance of \( N \, y(1) \).** From the above

\[
\text{Var}(N \, y(1) \mid L) = M \, (N-1) \, s_{all \, yy(j), \, j \leq M}^2
\]

Using this in a decomposition of unconditional variance of

\[
\text{Var}(N \, y(1)) = \text{Var}( E(N \, y(1) \mid L)) + E \, \text{Var}(N \, y(1) \mid L)
\]

\[
= \text{Var}(\tau) + E \, M \, (N-1) \, s_{all \, yy(j), \, j \leq M}^2
\]

\[
= M \, \gamma + M \, (N-1) \, \gamma = M \, N \, \gamma
\]

But \( \text{Var}(N \, y(1)) = E(Ny(1) \, - E(N \, y(1)))^2 = E(Ny(1) \, - E\tau)^2 \) and we wish to predict \( \tau \) not estimate its unconditional expectation.

**Expected (model) squared error of prediction of \( \tau \).**

\[
E(Ny(1) \, - \tau)^2 = E \, E((Ny(1) \, - \tau)^2 \mid L)
\]

\[
= E \, \text{Var}(Ny(1) \mid L) \quad \text{since} \quad E(Ny(1) \mid L) = \tau
\]

\[
\quad \quad \quad \quad \quad = E \, M \, (N-1) \, s_{all \, yy(j), \, j \leq M}^2 \quad \text{squared unbiased}
\]

\[
\quad \quad \quad \quad \quad = M \, (N-1) \, \gamma .
\]

**Unconditional (model) zPredictionInterval for \( \tau \).** This is just like a zCI except it uses the root mean squared prediction error calculated just above.

\[
N \, y(1) \pm z \, \sqrt{M(N - 1) \, s_{yy(j), \, j \, \in \, PU(1)}^2}, \quad E \, s_{yy(j), \, j \, \in \, PU(1)}^2 = \gamma .
\]
**Conclusion.** Unconditional zPredictionInterval and \( L \)-conditional zCI both use the exact same estimator of \( \tau \) and z-interval. They operate exactly the same, making the same claims.

Their shared purpose is to estimate the total \( \tau \) of random \( yy \) scores dealt iid into the problem. However the unconditional model has to address this as a problem of prediction owing to the fact that \( \tau \) is itself random. This leads to a need to examine CLT for prediction, which we have not done.

In the conditional model, conditional on \( L = \) the unordered list of \( yy(j), j \leq M \) the population total \( \tau \) is a constant parameter and \( y(1) \) is the sample total of a SRS of \( m \) from a population having scores \( L \). The conditional model addresses our objective clearly, simplifying the calculations through conditioning on \( L \) relative to which \( \tau \) is constant, and using SRS formulas and CLT.

**Design unbiasedness and variance of estimator** \( N \ y(U) \). Denote by \( U \) a random integer in 1 to \( N \).

\[
E \ N \ y(U) = N \ (1/N) \ E(y(1) + .. + y(N)) = \tau \\
Var(N \ y(U)) = \sum_1^N (N \ y(i) - \tau)^2/N
\]

**Estimation of design variance of estimator** \( N \ y(U) \). Why (specifically) can we not use only design-probabilities (i.e. uniform random selection in the estimator \( N \ y(U) \)) to estimate (design) variance of \( N \ y(U) \)? The naive answer is we only have a SRS of 1 from \( y(i), i \leq N \). A simple proof can be made.