Ratio estimation with Stratified Random Sampling

When using ratio estimation with stratified random sampling, there are two different ways to produce estimates. One way is to perform ratio estimation separately in each stratum, and then combine them. This gives a **separate ratio estimator**. The second way is to compute estimators for μ_y and μ_x using estimators for stratified random sampling, and then use $\bar{y}_{st}/\bar{x}_{st}$ as a ratio estimator of μ_y/μ_x. This gives a **combined ratio estimator**.

In a situation with two strata (labeled A and B), the expression for estimating a mean using the separate ratio estimator is:

$$\hat{\mu}_{yRS} = \left(\frac{N_A}{N}\right)\left(\frac{\bar{y}_A}{\bar{x}_A}\right)(\mu_{xA}) + \left(\frac{N_B}{N}\right)\left(\frac{\bar{y}_B}{\bar{x}_B}\right)(\mu_{xB}),$$

with estimated variance:

$$\hat{V}(\hat{\mu}_{yRS}) = \left(\frac{N_A}{N}\right)^2 \left(\frac{N_A - n_A}{N_A}\right)\left(\frac{1}{n_A}\right)\frac{\sum_{i=1}^{n_A}(y_i - r_A x_i)^2}{n_A - 1}$$

$$+ \left(\frac{N_B}{N}\right)^2 \left(\frac{N_B - n_B}{N_B}\right)\left(\frac{1}{n_B}\right)\frac{\sum_{i=1}^{n_B}(y_i - r_B x_i)^2}{n_B - 1}.$$

In a situation with two strata (labeled A and B), the expression for estimating a mean using the combined ratio estimator is:

$$\hat{\mu}_{yRC} = \frac{\bar{y}_{st}}{\bar{x}_{st}}(\mu_x), \text{ or } \hat{\mu}_{yRC} = r_C(\mu_x), \text{ where } r_C = \frac{\bar{y}_{st}}{\bar{x}_{st}},$$

with estimated variance:

$$\hat{V}(\hat{\mu}_{yRC}) = \left(\frac{N_A}{N}\right)^2 \left(\frac{N_A - n_A}{N_A}\right)\left(\frac{1}{n_A}\right)\frac{\sum_{i=1}^{n_A}(y_i - r_C x_i)^2}{n_A - 1}$$

$$+ \left(\frac{N_B}{N}\right)^2 \left(\frac{N_B - n_B}{N_B}\right)\left(\frac{1}{n_B}\right)\frac{\sum_{i=1}^{n_B}(y_i - r_C x_i)^2}{n_B - 1}.$$

Which approach should we use? Generally, the concern with the separate ratio estimator is that with small sample sizes per stratum, the individual stratum variance estimates will be biased, and that bias is added across strata. Thus it is recommended to use the separate ratio estimator unless the stratum sizes are small, say $(n_i < 20)$, or if the within-stratum ratios are approximately equal. Estimates of population totals are obtained by multiplication by the population size N, giving $\hat{\tau}_{yRS} = N\hat{\mu}_{yRS}$ or $\hat{\tau}_{yRC} = N\hat{\mu}_{yRC}$.

Examples 6.7 and 6.8 illustrate these estimators for a situation with two strata: