BRIEF ARTICLE

THE AUTHOR

1. Optimality of the AHL Estimator

Here we consider the possibility of incorporating probabilities of matches in the AHL framework. The interesting result is that the minimum variance unbiased estimator depends on the ratio of the structural error in the model to the variance in the reduced form estimation for y, which also depends on the precision of the estimator $\hat{g}$. This is an interesting result because it connects to the Horwitz-Thompson estimator, and work on inverse propensity score weighting.

To simplify ideas, consider the case of estimating the mean μ of a random variable $X \sim F_X(\mu; \sigma^2)$ using two observations X_{i1} and X_{i2}. The first observation is drawn from the true distribution $F_X(\mu; \sigma^2)$ with probability π, and from the nuisance distribution $F_N(\kappa, \omega^2)$ with probability $1 - \pi$. One of the observations is drawn from the correct distribution, so that π of the time, X_{i1} is correct, and $1 - \pi$ of the time X_{i2} is correct.

Consider the estimator of the mean

$$\hat{\mu} = \frac{a_1}{n}X_{i1} + \frac{a_2}{n}X_{i2} - a_3\kappa$$

and

$$E[\hat{\mu}] = (a_1\pi + a_2(1 - \pi))\mu + (a_1(1 - \pi) + a_2\pi - a_3)\kappa$$

1

So for $\hat{\mu}$ to be unbiased, we need

(1)
$$a_1 \pi + a_2(1 - \pi) = 1$$

(2)
$$a_1(1 - \pi) + a_2 \pi = a_3$$

A simpler way to solve the problem is to consider estimating the mean with just X_{i1}:

$$\hat{\mu}_1 = b_1 X_1 + b_3 \kappa$$

$$E[\hat{\mu}] = b_1 \pi \mu + [b_1(1 - \pi) - b_3]\kappa$$

Unbiasedness therefore requires that $b_1 = \frac{1}{\pi}$ and $b_3 = \frac{1-\pi}{\pi}$. Similarly we construct

$$\hat{\mu}_2 = C_1 X_2 - C_2 \kappa$$

which gives the condition $C_2 = \frac{1}{1-\pi}$ and $C_3 = \frac{\pi}{1-\pi}$ for unbiasedness. Connection to Horwitz-Thompson?

Any unbiased estimator of the form CITE $\hat{\mu}$ can be written as a linear combination of $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$\hat{\mu} = d_1 \hat{\mu}_1 + d_2 \hat{\mu}_2$$

The variance of $\hat{\mu}$ depends on:

$$\text{Var}\,(\hat{\mu}_1) = \frac{\text{Var}(X_1)}{\pi^2} = \frac{1}{\pi^2}\left(\pi\sigma^2 + (1 - \pi)\omega^2 + \pi(1 - \pi)(\mu - \kappa)^2\right)$$

$$\text{Var}\,(\hat{\mu}_2) = \frac{\text{Var}\,(X_2)}{(1 - \pi)^2} = \frac{1}{(1 - \pi)^2}\left((1 - \pi)\sigma^2 + \pi\omega^2 + \pi(1 - \pi)(\mu - \kappa)^2\right)$$

This follows from the law of total variance, with the random variable $D = 1$ if X_1 is drawn from the correct distribution and $D = 0$ otherwise.

$$\text{Var}\,(X_1) = E[\text{Var}\,(X_1|D)] + \text{Var}\,(E[X_1|D])$$

$$= P(D=1)\sigma^2 + P(D=0)\omega^2 + \text{Var}\,(()\,\mu D + \kappa(1-D))$$

$$= \pi\sigma^2 + (1-\pi)\omega^2 + \pi(1-\pi)(\mu-\kappa)^2$$

and similarly

$$\text{Var}\,(X_2) = (1-\pi)\sigma^2 + \pi\omega^2 + \pi(1-\pi)(\mu-\kappa)^2$$

The important observation is that the variances are functions parameterized by $(\sigma^2, \omega^2, (\mu-\kappa)^2)$, and we can write:

$$\text{Var}\,(\hat{\mu}_1) = g(\sigma^2, \omega^2, (\mu-\kappa)^2, \pi)$$

$$\text{Var}\,(\hat{\mu}_2) = g(\sigma^2, \omega^2, (\mu-\kappa)^2, 1-\pi)$$

They are exactly equal when $\pi = 1 - \pi$; otherwise they are quadratic as in the plot below. But they are exactly symmetric, so that the optimal estimator will always assign $d_1 = 1, d_2 = 0$ or $d_1 = 0, d_2 = 1$ depending on whether $\pi > 1 - \pi$. However, if π is not estimated precisely, the resulting variance explodes. The most conservative way is by treating $\pi = 0.5$, which gives the estimate of the mean! Which is the AHL estimator.