`Survey` package: Ratio estimation

Week 4

Stat 260, St. Clair

Ratio estimators

$$\hat{B} = \frac{\bar{y}}{\bar{x}}$$

- Estimate $\hat{B}$ and SE

```
> est.B <- svyratio(~y, ~x, design)
```

- Estimate $\hat{t}_{y,r}$ and SE   total

```
> predict(est.B, tx)    # tx = KNOWN POPULATION total for x
```

- Estimate $\hat{\bar{y}}_r$ and SE   mean

```
> predict(est.B, mnx)    # tx = KNOWN POPULATION mean for x
```

$$\hat{t}_{y,r} = \hat{B}\, t_x$$

$$\hat{\bar{y}}_r = \hat{B}\, \bar{x}_u$$

Lohr Examples 4.2 and 4.3

We have a SRS of 300 counties taken from the population of 3078 counties.

```
> library(survey)
> library(SDaA)
> agsrs$n<- nrow(agsrs)
> agsrs$N<- 3078
> agsrs$wts<- agsrs$N/agsrs$n
> design.srs<- svydesign(id= ~1,
+                        fpc= ~N,
+                        weights= ~wts,
+                        data= agsrs)
```

not new (ch. 2)

Lohr Examples 4.2 and 4.3

Estimate the ratio of total 1992 acres devoted to farming (y) to the total from 1987 (x)

$$B = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i}$$

```
> ratio<-svyratio(~acres92, ~acres87, design.srs)
> ratio   #estimate/SE for B
Ratio estimator: svyratio.survey.design2(~acres92, ~acres87, design.s
Ratios=
          acres87
acres92 0.9865652
SEs=
            acres87
acres92 0.005750473
> confint(ratio, df=degf(design.srs))  # CI for B
                   2.5 %    97.5 %
acres92/acres87 0.9752487 0.9978818
```

$\hat{B}$

$SE(\hat{B})$

Lohr Examples 4.2 and 4.3

Known: total number of acres devoted to farming in 1987 was
$t_x = 964,470,625$

Estimate the total number of farming acres in 1992 using a **ratio** estimate:

```
> ratio<-svyratio(~acres92, ~acres87, design.srs)
> tx <- 964470625     # pop. total from 1987 (example 4.2)
> toty <- predict(ratio , tx)    # ratio estimate of 1992 total
> toty
$total
          acres87
acres92 951513191

$se
        acres87
acres92 5546162
> confint(ratio, df=degf(design.srs)) *tx    # ratio est. CI for 1992
                    2.5 %    97.5 %
acres92/acres87 940598734 962427648
```

Lohr Examples 4.2 and 4.3

Compare the ratio estimate of total to the SRS estimate:

```
> toty
$total
            acres87
acres92 951513191

$se
          acres87
acres92 5546162
> svytotal(~acres92,design.srs)    # compare to usual SRS est. of 199
              total        SE
acres92 916927110 58169381
```

– ratio est $\hat{t}_{y,r}$

\# much more precise

$\hat{t}_{y,SRS}$

Lohr Examples 4.2 and 4.3

Known: mean number of acres per county devoted to farming in 1987 was $964,470,625/3078$ N

Estimate the mean number of farming acres per county in 1992 using a **ratio** estimate:

```
> mnx<- tx/3078    # pop. mean from 1987
> mny<-predict(ratio , mnx)  # ratio estimate of 1992 mean
> mny
$total
         acres87
acres92 309133.6

$se
         acres87
acres92 1801.872
> confint(ratio, df=degf(design.srs)) *mnx   # ratio CI of 1992 mean
                   2.5 %   97.5 %
acres92/acres87 305587.6 312679.5
> svymean(~acres92, design.srs)    # usual SRS estimate of 1992 mean
          mean    SE
acres92 297897 18898
```

Lohr Examples 4.2 and 4.3

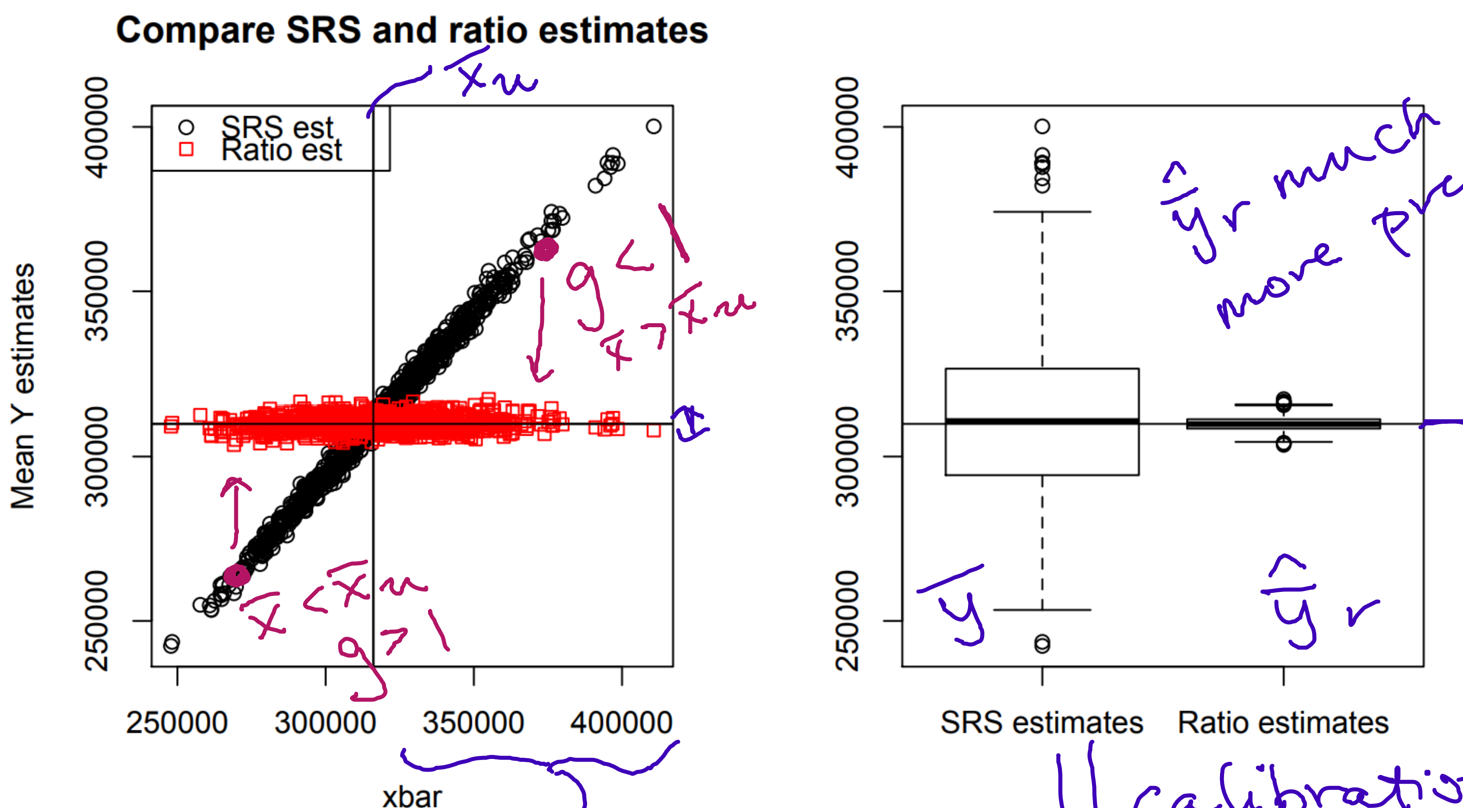
Simulation: lots of SRS of $n=300$ from pop. For each: $\bar{y}$, $\bar{x}$, $\hat{\bar{y}}_r$

Why are the ratio estimates of 1992 total/mean more precise that SRS?

y $\qquad$ $x = 1987$ acres

$\bar{y}$ = black

$\hat{\bar{y}}_r$ = red



$\bar{y}_U$

$\bar{x}$ $\qquad$ $\bar{x}$ above pop. mean $\bar{x}_U$

calibration weight

$$g = \frac{t_x}{\hat{t}_x} = \frac{\bar{x}_U}{\bar{x}}$$