Api

The Academic Performance Index is computed for all California schools based on standardised testing of students. The data sets contain information for all schools with at least 100 students and for various probability samples of the data.

stype

: Elementary/Middle/High School

api00

: API in 2000

api99

: API in 1999

meals

: Percentage of students eligible for subsidized meals

ell

: 'English Language Learners' (percent)

enroll

: number of students enrolled

full

: percent fully qualified teachers

sch.wide

: Met school-wide growth target?

comp.imp

: Met Comparable Improvement target

awards

: Eligible for awards program

meals

: Percentage of students eligible for subsidized meals

The other data sets contain additional variables `pw` for sampling weights and `fpc` to compute finite population corrections to variance.

Details

`apipop` is the entire population,

`apisrs` is a simple random sample,

`apiclus1` is a cluster sample of school districts,

`apistrat` is a sample stratified by `stype`,

and `apiclus2` is a two-stage cluster sample of schools within districts.

The sampling weights in `apiclus1` are incorrect (the weight should be 757/15) but are as obtained from UCLA.

```
> library(survey)
> data(api)###  6194 observaciones con 37 variables
> mean(apipop$api00)
[1] 664.7126
> sum(apipop$enroll, na.rm=TRUE)
[1] 3811472
>
> #stratified sample
> dstrat<-svydesign(id=~1,strata=~stype, weights=~pw, data=apistrat, fpc=~fpc)
> summary(dstrat)
Stratified Independent Sampling design
svydesign(id = ~1, strata = ~stype, weights = ~pw, data = apistrat,
    fpc = ~fpc)
Probabilities:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.02262 0.02262 0.03587 0.04014 0.05339 0.06623
Stratum Sizes:
             E  H  M
obs        100 50 50
design.PSU 100 50 50
actual.PSU 100 50 50
Population stratum sizes (PSUs):
   E    H    M
4421  755 1018
Data variables:
 [1] "cds"      "stype"    "name"     "sname"    "snum"     "dname"    "dnum"     "cname"    "c
[11] "pcttest"  "api00"    "api99"    "target"   "growth"   "sch.wide" "comp.imp" "both"     "a
[21] "ell"      "yr.rnd"   "mobility" "acs.k3"   "acs.46"   "acs.core" "pct.resp" "not.hsg"  "h
ol"
[31] "col.grad" "grad.sch" "avg.ed"   "full"     "emer"     "enroll"   "api.stu"  "pw"       "f
> svymean(~api00, dstrat)
        mean     SE
api00 662.29 9.4089
> svytotal(~enroll, dstrat, na.rm=TRUE)
          total     SE
enroll 3687178 114642
```

Definir el diseño del muestreo

n=100+50+50

Como correr los modelos glm

CASO 1. Una regresión lineal

```
> (svyglm(api00~ell+meals+full+api99,design=dstrat))
Stratified Independent Sampling design
svydesign(id = ~1, strata = ~stype, weights = ~pw, data = apistrat,
    fpc = ~fpc)

Call:  svyglm(formula = api00 ~ ell + meals + full + api99, design = dstrat)

Coefficients:
(Intercept)          ell        meals         full        api99
    19.2397      -0.2653       0.4022       0.1906       0.9742

Degrees of Freedom: 199 Total (i.e. Null);  193 Residual
Null Deviance:      3023000
Residual Deviance: 137300      AIC: 1907
> (glm(api00~ell+meals+full+api99,data=apistrat))## sin tomar en cuenta el  diseño

Call:  glm(formula = api00 ~ ell + meals + full + api99, data = apistrat)

Coefficients:
(Intercept)          ell        meals         full        api99
   -11.5421      -0.2271       0.4942       0.2365       1.0025

Degrees of Freedom: 199 Total (i.e. Null);  195 Residual
Null Deviance:      2912000
```

n=200. gl=n-p-2=200-(5+2)

el 2 tiene que ver con número de estratos menos 1

n=200. gl=n-p=200-(5)

```
Residual Deviance: 137600      AIC: 1886
```

CASO 2. Una modelo con respuesta binaria

```
> (svyglm(I(sch.wide=="Yes")~comp.imp+meals+awards,design=dstrat,family=quasibinomial()))
Stratified Independent Sampling design
svydesign(id = ~1, strata = ~stype, weights = ~pw, data = apistrat,
    fpc = ~fpc)

Call:  svyglm(formula = I(sch.wide == "Yes") ~ comp.imp + meals + awards,
    design = dstrat, family = quasibinomial())

Coefficients:
(Intercept)  comp.impYes        meals    awardsYes
  1.676e-01   -2.085e+01    6.021e-04    4.130e+01

Degrees of Freedom: 199 Total (i.e. Null);  194 Residual
Null Deviance:      183.7
Residual Deviance: 94.81       AIC: NA
```

Recomiendan usar así para evitar ciertos *warnings*

n=200. gl=n-p-2=200-(6)

```
> (glm(I(sch.wide=="Yes")~comp.imp+meals+awards,data=apistrat, family=binomial))## sin tomar en
ño

Call:  glm(formula = I(sch.wide == "Yes") ~ comp.imp + meals + awards,
    family = binomial, data = apistrat)

Coefficients:
(Intercept)  comp.impYes        meals    awardsYes
   0.085155   -20.418585    -0.005204    41.145161

Degrees of Freedom: 199 Total (i.e. Null);  196 Residual
Null Deviance:      220.4
Residual Deviance: 115.6       AIC: 123.6
```

n=200. gl=n-p=200-(4)