Remarks and examples Reference Also see

Remarks and examples

Stata's ml command can fit maximum likelihood-based models for survey data. Many ml-based estimators can now be modified to handle one or more stages of clustering, stratification, sampling weights, finite population correction, calibration, poststratification, and subpopulation estimation. See [R] ml for details.

See [P] **program properties** for a discussion of the programming requirements for an estimation command to work with the svy prefix. See Pitblado, Poi, and Gould (2024) for examples of community-contributed estimation commands that support the svy prefix.

Example 1: User-written survey regression

The ml command requires a program that computes likelihood values to perform maximum likelihood. Here is a likelihood evaluator used in Pitblado, Poi, and Gould (2024) to fit linear regression models using likelihood from the normal distribution.

```
program mynormal_lf
    version 19.5 // (or version 19 if you do not have StataNow)
    args lnf mu lnsigma
    quietly replace 'lnf' = ln(normalden($ML_y1,'mu',exp('lnsigma')))
end
```

Here we fit a survey regression model using a multistage survey dataset with ml and the above likelihood evaluator.

```
. use https://www.stata-press.com/data/r19/multistage
. svyset county [pw=sampwgt], strata(state) fpc(ncounties) || school,
> fpc(nschools)
Sampling weights: sampwgt
            VCE: linearized
    Single unit: missing
        Strata 1: state
 Sampling unit 1: county
          FPC 1: ncounties
        Strata 2: <one>
 Sampling unit 2: school
          FPC 2: nschools
. ml model lf mynormal lf (mu: weight = height) /lnsigma, svy
. ml max
Initial:
              Log pseudolikelihood =
                                         -<inf>
                                                 (could not be evaluated)
Feasible:
              Log pseudolikelihood = -7.301e+08
Rescale:
             Log pseudolikelihood =
                                     -51944380
Rescale eq:
             Log pseudolikelihood = -47565331
Iteration 0:
             Log pseudolikelihood = -47565331
Iteration 1: Log pseudolikelihood = -41225560 (not concave)
Iteration 2:
             Log pseudolikelihood = -41221017 (not concave)
Iteration 3: Log pseudolikelihood = -41170786
                                                (not concave)
Iteration 4: Log pseudolikelihood = -41151252
                                                (not concave)
Iteration 5: Log pseudolikelihood = -41132554 (not concave)
Iteration 6: Log pseudolikelihood = -41107910
                                                (not concave)
Iteration 7: Log pseudolikelihood = -41077230
Iteration 8: Log pseudolikelihood = -39815341
                                                (backed up)
Iteration 9: Log pseudolikelihood = -38344045
Iteration 10: Log pseudolikelihood = -38328784
Iteration 11: Log pseudolikelihood =
                                      -38328739
Iteration 12: Log pseudolikelihood = -38328739
Number of strata = 50
                                                   Number of obs
                                                                   =
                                                                         4,071
Number of PSUs
               = 100
                                                   Population size = 8,000,000
                                                   Design df
                                                                   =
                                                                            50
                                                   F(1, 50)
                                                                   =
                                                                        593.99
                                                   Prob > F
                                                                        0.0000
                                                                   =
                           Linearized
      weight
               Coefficient std. err.
                                                P>|t|
                                                          [95% conf. interval]
                                           t
      height
                 .7163115
                            .0293908
                                        24.37
                                                0.000
                                                          .6572784
                                                                       .7753447
       _cons
                -149.6183
                            12.57265
                                       -11.90
                                                0.000
                                                         -174.8712
                                                                     -124.3654
                3.372154
                            .0180777
                                       186.54
                                                0.000
                                                          3.335844
                                                                      3.408464
    /lnsigma
```

Reference

Pitblado, J. S., B. P. Poi, and W. W. Gould. 2024. *Maximum Likelihood Estimation with Stata*. 5th ed. College Station, TX: Stata Press.

Also see

- [SVY] Survey Introduction to survey commands
- [P] program properties Properties of user-defined programs
- [R] Maximize Details of iterative maximization
- [R] ml Maximum likelihood estimation

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow and NetCourseNow are trademarks of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.



For suggested citations, see the FAQ on citing Stata documentation.