

**ml for svy** — Maximum pseudolikelihood estimation for survey data

[Remarks and examples](#)

[Reference](#)

[Also see](#)

## Remarks and examples

stata.com

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, 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 [Gould, Pitblado, and Poi \(2010\)](#) for examples of user-written 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 [Gould, Pitblado, and Poi \(2010\)](#) to fit linear regression models using likelihood from the normal distribution.

```

program mynormal_lf
    version 14.1
    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 http://www.stata-press.com/data/r14/multistage
. svyset county [pw=sampwgt], strata(state) fpc(ncounties) || school,
> fpc(nschools)
      pweight: sampwgt
          VCE: linearized
Single unit: missing
  Strata 1: state
           SU 1: county
           FPC 1: ncounties
  Strata 2: <one>
           SU 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 = -41226725 (not concave)
Iteration 2:  log pseudolikelihood = -41221650 (not concave)
Iteration 3:  log pseudolikelihood = -41176159 (not concave)
Iteration 4:  log pseudolikelihood = -41154139 (not concave)
Iteration 5:  log pseudolikelihood = -41052368
Iteration 6:  log pseudolikelihood = -39379181 (backed up)
Iteration 7:  log pseudolikelihood = -38333242
Iteration 8:  log pseudolikelihood = -38328742
Iteration 9:  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
```

	weight	Linearized			P> t	[95% Conf. Interval]	
		Coef.	Std. Err.	t			
mu							
	height	.716311	.0293908	24.37	0.000	.6572778	.7753442
	_cons	-149.6181	12.57266	-11.90	0.000	-174.871	-124.3652
lnsigma							
	_cons	3.372153	.0180777	186.54	0.000	3.335843	3.408464



## Reference

Gould, W. W., J. S. Pitblado, and B. P. Poi. 2010. *Maximum Likelihood Estimation with Stata*. 4th ed. College Station, TX: Stata Press.

## Also see

- [P] [program properties](#) — Properties of user-defined programs
- [R] [maximize](#) — Details of iterative maximization
- [R] [ml](#) — Maximum likelihood estimation
- [SVY] [survey](#) — Introduction to survey commands