mestreg — Multilevel mixed-effects parametric survival models

Description

mestreg fits a mixed-effects parametric survival-time model. The conditional distribution of the response given the random effects is assumed to be an exponential, Weibull, lognormal, loglogistic, or gamma distribution. mestreg can be used with single- or multiple-record st data.

Quick start

Without weights

Two-level Weibull survival model with covariates x1 and x2 and random intercepts by lev2 using stset data

```
mestreg x1 x2 || lev2:, distribution(weibull)
```

Mixed-effects model adding random coefficients for x1

```
mestreg x1 x2 || lev2:x1, distribution(weibull)
```

Three-level random-intercept model with lev2 nested within lev3

```
mestreg x1 x2 || lev3: || lev2:, distribution(weibull)
```

With weights

Two-level Weibull survival model with covariates x1 and x2, random intercepts by lev2, and observation-level frequency weights wvar1 using stset data

```
mestreg x1 x2 [fweight=wvar1] || lev2:, distribution(weibull)
```

Two-level random-intercept model from a two-stage sampling design with PSUs identified by psu using PSU-level and observation-level sampling weights wvar2 and wvar1

```
mestreg x1 x2 [pweight=wvar1] || psu:, pweight(wvar2)
```

Same as above, but svyset the data first

```
svyset psu, weight(wvar2) || _n, weight(wvar1)
svy: mestreg x1 x2 || psu:, distribution(weibull)
```

Note: Any supported parametric survival distribution may be specified in place of weibull above.

Menu

Statistics > Multilevel mixed-effects models > Parametric survival regression

Syntax

`mestreg` *fe_equation* [|| *re_equation*] [|| *re_equation* ...] ,

  `distribution(`*distname*`)` [*options*]

where the syntax of *fe_equation* is

  [*indepvars*] [*if*] [*in*] [*weight*] [, *fe_options*]

and the syntax of *re_equation* is one of the following:

 for random coefficients and intercepts

  levelvar: [*varlist*] [, *re_options*]

 for random effects among the values of a factor variable in a crossed-effects model

  levelvar: `R.`*varname*

levelvar is a variable identifying the group structure for the random effects at that level or is `_all` representing one group comprising all observations.

fe_options	Description
Model	
`noconstant`	suppress constant term from the fixed-effects equation
`offset(`*varname*`)`	include *varname* in model with coefficient constrained to 1

re_options	Description
Model	
`covariance(`*vartype*`)`	variance–covariance structure of the random effects
`noconstant`	suppress constant term from the random-effects equation
`fweight(`*varname*`)`	frequency weights at higher levels
`iweight(`*varname*`)`	importance weights at higher levels
`pweight(`*varname*`)`	sampling weights at higher levels

options	Description
Model	
* distribution(*distname*)	specify survival distribution
time	use accelerated failure-time metric
constraints(*constraints*)	apply specified linear constraints
SE/Robust	
vce(*vcetype*)	*vcetype* may be oim, opg, robust, or cluster *clustvar*
Reporting	
level(#)	set confidence level; default is level(95)
nohr	do not report hazard ratios
tratio	report time ratios
noshow	do not show st setting information
nocnsreport	do not display constraints
notable	suppress coefficient table
noheader	suppress output header
nogroup	suppress table summarizing groups
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Integration	
intmethod(*intmethod*)	integration method
intpoints(#)	set the number of integration (quadrature) points for all levels; default is intpoints(7)
Maximization	
maximize_options	control the maximization process; seldom used
startvalues(*svmethod*)	method for obtaining starting values
startgrid[(*gridspec*)]	perform a grid search to improve starting values
noestimate	do not fit the model; show starting values instead
dnumerical	use numerical derivative techniques
collinear	keep collinear variables
coeflegend	display legend instead of statistics

*distribution(*distname*) is required.

vartype	Description
independent	one unique variance parameter per random effect and all covariances 0; the default unless the R. notation is used
exchangeable	equal variances for random effects and one common pairwise covariance
identity	equal variances for random effects and all covariances 0; the default if the R. notation is used
unstructured	all variances and covariances to be distinctly estimated
fixed(*matname*)	user-selected variances and covariances constrained to specified values; the remaining variances and covariances unrestricted
pattern(*matname*)	user-selected variances and covariances constrained to be equal; the remaining variances and covariances unrestricted

distname	Description
`exponential`	exponential survival distribution
`loglogistic`	loglogistic survival distribution
`llogistic`	synonym for `loglogistic`
`weibull`	Weibull survival distribution
`lognormal`	lognormal survival distribution
`lnormal`	synonym for `lognormal`
`gamma`	gamma survival distribution

intmethod	Description
`mvaghermite`	mean–variance adaptive Gauss–Hermite quadrature; the default unless a crossed random-effects model is fit
`mcaghermite`	mode-curvature adaptive Gauss–Hermite quadrature
`ghermite`	nonadaptive Gauss–Hermite quadrature
`laplace`	Laplacian approximation; the default for crossed random-effects models

You must `stset` your data before using `mestreg`; see [ST] **stset**.

indepvars and *varlist* may contain factor variables; see [U] **11.4.3 Factor variables**.

`bayes`, `by`, `collect`, and `svy` are allowed; see [U] **11.1.10 Prefix commands**. For more details, see [BAYES] **bayes: mestreg**.

`vce()` and weights are not allowed with the `svy` prefix; see [SVY] **svy**.

`fweights`, `iweights`, and `pweights` are allowed; see [U] **11.1.6 weight**. Only one type of weight may be specified. Weights are not supported under the Laplacian approximation or for crossed models.

`startvalues()`, `startgrid`, `noestimate`, `dnumerical`, `collinear`, and `coeflegend` do not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Options

Model

`noconstant` suppresses the constant (intercept) term and may be specified for the fixed-effects equation and for any of or all the random-effects equations.

`offset(`*varname*`)` specifies that *varname* be included in the fixed-effects portion of the model with the coefficient constrained to be 1.

`covariance(`*vartype*`)` specifies the structure of the covariance matrix for the random effects and may be specified for each random-effects equation. *vartype* is one of the following: `independent`, `exchangeable`, `identity`, `unstructured`, `fixed(`*matname*`)`, or `pattern(`*matname*`)`.

`covariance(independent)` covariance structure allows for a distinct variance for each random effect within a random-effects equation and assumes that all covariances are 0. The default is `covariance(independent)` unless a crossed random-effects model is fit, in which case the default is `covariance(identity)`.

`covariance(exchangeable)` structure specifies one common variance for all random effects and one common pairwise covariance.

`covariance(identity)` is short for "multiple of the identity"; that is, all variances are equal and all covariances are 0.

`covariance(unstructured)` allows for all variances and covariances to be distinct. If an equation consists of p random-effects terms, the unstructured covariance matrix will have $p(p+1)/2$ unique parameters.

`covariance(fixed(`*matname*`))` and `covariance(pattern(`*matname*`))` covariance structures provide a convenient way to impose constraints on variances and covariances of random effects. Each specification requires a *matname* that defines the restrictions placed on variances and covariances. Only elements in the lower triangle of *matname* are used, and row and column names of *matname* are ignored. A missing value in *matname* means that a given element is unrestricted. In a `fixed(`*matname*`)` covariance structure, (co)variance (i, j) is constrained to equal the value specified in the i, jth entry of *matname*. In a `pattern(`*matname*`)` covariance structure, (co)variances (i, j) and (k, l) are constrained to be equal if *matname*$[i, j]$ = *matname*$[k, l]$.

`fweight(`*varname*`)` specifies frequency weights at higher levels in a multilevel model, whereas frequency weights at the first level (the observation level) are specified in the usual manner, for example, `[fw=`*fwtvar1*`]`. *varname* can be any valid Stata variable name, and you can specify `fweight()` at levels two and higher of a multilevel model. For example, in the two-level model

```
. mecmd fixed_portion [fw = wt1] || school: ... , fweight(wt2) ...
```

the variable `wt1` would hold the first-level (the observation-level) frequency weights, and `wt2` would hold the second-level (the school-level) frequency weights.

`iweight(`*varname*`)` specifies importance weights at higher levels in a multilevel model, whereas importance weights at the first level (the observation level) are specified in the usual manner, for example, `[iw=`*iwtvar1*`]`. *varname* can be any valid Stata variable name, and you can specify `iweight()` at levels two and higher of a multilevel model. For example, in the two-level model

```
. mecmd fixed_portion [iw = wt1] || school: ... , iweight(wt2) ...
```

the variable `wt1` would hold the first-level (the observation-level) importance weights, and `wt2` would hold the second-level (the school-level) importance weights.

`pweight(`*varname*`)` specifies sampling weights at higher levels in a multilevel model, whereas sampling weights at the first level (the observation level) are specified in the usual manner, for example, `[pw=`*pwtvar1*`]`. *varname* can be any valid Stata variable name, and you can specify `pweight()` at levels two and higher of a multilevel model. For example, in the two-level model

```
. mecmd fixed_portion [pw = wt1] || school: ... , pweight(wt2) ...
```

variable `wt1` would hold the first-level (the observation-level) sampling weights, and `wt2` would hold the second-level (the school-level) sampling weights.

`distribution(`*distname*`)` specifies the survival model to be fit. *distname* is one of the following: `exponential`, `loglogistic`, `llogistic`, `weibull`, `lognormal`, `lnormal`, or `gamma`. This option is required.

`time` specifies that the model be fit in the accelerated failure-time metric rather than in the log relative-hazard metric. This option is valid only for the exponential and Weibull models because these are the only models that have both a proportional-hazards and an accelerated failure-time parameterization. Regardless of metric, the likelihood function is the same, and models are equally appropriate in either metric; it is just a matter of changing interpretation.

`time` must be specified at estimation.

`constraints(`*constraints*`)`; see [R] **Estimation options**.

SE/Robust

vce(*vcetype*) specifies the type of standard error reported, which includes types that are derived from asymptotic theory (oim, opg), that are robust to some kinds of misspecification (robust), and that allow for intragroup correlation (cluster *clustvar*); see [R] ***vce_option***. If vce(robust) is specified, robust variances are clustered at the highest level in the multilevel model.

Reporting

level(#); see [R] **Estimation options**.

nohr, which may be specified at estimation or upon redisplaying results, specifies that coefficients rather than exponentiated coefficients be displayed, that is, that coefficients rather than hazard ratios be displayed. This option affects only how coefficients are displayed, not how they are estimated.

This option is valid only for the exponential and Weibull models because they have a natural proportional-hazards parameterization. These two models, by default, report hazards ratios (exponentiated coefficients).

tratio specifies that exponentiated coefficients, which are interpreted as time ratios, be displayed. tratio is appropriate only for the loglogistic, lognormal, and gamma models or for the exponential and Weibull models when fit in the accelerated failure-time metric.

tratio may be specified at estimation or upon replay.

noshow prevents mestreg from showing the key st variables. This option is rarely used because most users type stset, show or stset, noshow to set once and for all whether they want to see these variables mentioned at the top of the output of every st command; see [ST] **stset**.

nocnsreport; see [R] **Estimation options**.

notable suppresses the estimation table, either at estimation or upon replay.

noheader suppresses the output header, either at estimation or upon replay.

nogroup suppresses the display of group summary information (number of groups, average group size, minimum, and maximum) from the output header.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlabel, fvwrap(#), fvwrapon(*style*), cformat(%*fmt*), pformat(%*fmt*), sformat(%*fmt*), and nolstretch; see [R] **Estimation options**.

Integration

`intmethod(`*intmethod*`)` specifies the integration method to be used for the random-effects model. `mvaghermite` performs mean–variance adaptive Gauss–Hermite quadrature; `mcaghermite` performs mode-curvature adaptive Gauss–Hermite quadrature; `ghermite` performs nonadaptive Gauss–Hermite quadrature; and `laplace` performs the Laplacian approximation, equivalent to mode-curvature adaptive Gaussian quadrature with one integration point.

The default integration method is `mvaghermite` unless a crossed random-effects model is fit, in which case the default integration method is `laplace`. The Laplacian approximation has been known to produce biased parameter estimates; however, the bias tends to be more prominent in the estimates of the variance components rather than in the estimates of the fixed effects.

For crossed random-effects models, estimation with more than one quadrature point may be prohibitively intensive even for a small number of levels. For this reason, the integration method defaults to the Laplacian approximation. You may override this behavior by specifying a different integration method.

`intpoints(#)` sets the number of integration points for quadrature. The default is `intpoints(7)`, which means that seven quadrature points are used for each level of random effects. This option is not allowed with `intmethod(laplace)`.

The more integration points, the more accurate the approximation to the log likelihood. However, computation time increases as a function of the number of quadrature points raised to a power equaling the dimension of the random-effects specification. In crossed random-effects models and in models with many levels or many random coefficients, this increase can be substantial.

Maximization

maximize_options: `difficult`, `technique(`*algorithm_spec*`)`, `iterate(#)`, `[no]log`, `trace`, `gradient`, `showstep`, `hessian`, `showtolerance`, `tolerance(#)`, `ltolerance(#)`, `nrtolerance(#)`, `nonrtolerance`, and `from(`*init_specs*`)`; see [R] **Maximize**. Those that require special mention for `mestreg` are listed below.

`from()` accepts a properly labeled vector of initial values or a list of coefficient names with values. A list of values is not allowed.

The following options are available with `mestreg` but are not shown in the dialog box:

`startvalues(`*svmethod*`)`, `startgrid[(`*gridspec*`)]`, `noestimate`, and `dnumerical`; see [ME] **meglm**.

`collinear`, `coeflegend`; see [R] **Estimation options**.

Remarks and examples

For a general introduction to `me` commands, see [ME] **me**.

Remarks are presented under the following headings:

Introduction

Mixed-effects survival models contain both fixed effects and random effects. In longitudinal data and panel data, random effects are useful for modeling intracluster correlation; that is, observations in the same cluster are correlated because they share common cluster-level random effects.

mestreg allows for many levels of random effects. However, for simplicity, we now consider two-level models, where we have a series of M independent clusters and a set of random effects $\mathbf{u}_j$ corresponding to those clusters. Two often-used models for adjusting survivor functions for the effects of covariates are the accelerated failure-time (AFT) model and the multiplicative or proportional hazards (PH) model.

In the AFT model, the natural logarithm of the survival time, $\log t$, is expressed as a linear function of the covariates; when we incorporate random-effects, this yields the model

$$\log t_{ji} = \mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j + v_{ji}$$

for $j = 1, \ldots, M$ clusters, with cluster j consisting of $i = 1, \ldots, n_j$ observations. The $1 \times p$ row vector $\mathbf{x}_{ji}$ contains the covariates for the fixed effects, with regression coefficients (fixed effects) $\boldsymbol{\beta}$.

The $1 \times q$ vector $\mathbf{z}_{ji}$ contains the covariates corresponding to the random effects and can be used to represent both random intercepts and random coefficients. For example, in a random-intercept model, $\mathbf{z}_{ji}$ is simply the scalar 1. The random effects $\mathbf{u}_j$ are M realizations from a multivariate normal distribution with mean $\mathbf{0}$ and $q \times q$ variance matrix $\boldsymbol{\Sigma}$. The random effects are not directly estimated as model parameters but are instead summarized according to the unique elements of $\boldsymbol{\Sigma}$, known as variance components.

Finally, v_{ji} are the observation-level errors with density $\varphi(\cdot)$. The distributional form of the error term determines the regression model. Five regression models are implemented in mestreg using the AFT parameterization: exponential, gamma, loglogistic, lognormal, and Weibull. The lognormal regression model is obtained by letting $\varphi(\cdot)$ be the normal density. Similarly, by letting $\varphi(\cdot)$ be the logistic density, one obtains the loglogistic regression. Setting $\varphi(\cdot)$ equal to the extreme-value density yields the exponential and the Weibull regression models.

In the PH models fit by mestreg, the covariates have a multiplicative effect on the hazard function

$$h(t_{ji}) = h_0(t_{ji}) \exp(\mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j)$$

for some baseline hazard function $h_0(t)$. For the mestreg command, $h_0(t)$ is assumed to be parametric. The exponential and Weibull models are implemented in mestreg for the PH parameterization. These two models are implemented using both the AFT and PH parameterizations.

mestreg is suitable only for data that have been stset. By using stset on your data, you define the variables _t0, _t, and _d, which serve as the trivariate response variable (t_0, t, d). Each response corresponds to a period under observation, $(t_0, t]$, resulting in either failure $(d = 1)$ or right-censoring $(d = 0)$ at time t.

mestreg does not allow delayed entry or gaps. However, mestreg is appropriate for data exhibiting multiple records per subject and time-varying covariates. mestreg requires subjects to be nested within clusters.

stset weights are not used; instead, weights must be specified at estimation. Weights are not allowed with crossed models or the Laplacian approximation. See *Survey estimation* in *Methods and formulas* for details.

Two-level models

▷ Example 1: Two-level random-intercept PH model

In example 11 of [ST] **streg**, we fit a Weibull model with an inverse-Gaussian shared frailty to the recurrence times for catheter-insertion point infection for 38 kidney dialysis patients. In this example, the subjects are the catheter insertions, not the patients themselves. This is a function of how the data were recorded—the onset of risk occurs at the time the catheter is inserted and not, say, at the time of admission of the patient into the study. Thus we have two subjects (insertions) within each group (patient). Each catheter insertion results in either infection (`infect==1`) or right-censoring (`infect==0`). The `stset` results are shown below.

```
. use https://www.stata-press.com/data/r19/catheter
(Kidney data, McGilchrist and Aisbett, Biometrics, 1991)

. stset
-> stset time, failure(infect)

Survival-time data settings

         Failure event: infect!=0 & infect<.
Observed time interval: (0, time]
     Exit on or before: failure

--------------------------------------------------------------------------
         76  total observations
          0  exclusions
--------------------------------------------------------------------------
         76  observations remaining, representing
         58  failures in single-record/single-failure data
      7,424  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =       562
```

While it is reasonable to assume independence of patients, we would not want to assume that recurrence times within each patient are independent. The model used in [ST] **streg** allowed us to model the correlation by assuming that it was the result of a latent patient-level effect, or frailty.

The random-effects approach used by `mestreg` is more flexible because it allows you to experiment with several levels of random effects, including random coefficients, or both. You can then choose the model that best suits your data. Here we use `mestreg` to fit a random-effects Weibull model with normally distributed random effects. This model can be viewed as a shared frailty model with lognormal frailty.

```
. mestreg age female || patient:, distribution(weibull)
        Failure _d: infect
  Analysis time _t: time
Fitting fixed-effects model:
Iteration 0:  Log likelihood = -1700989.9
Iteration 1:  Log likelihood =  -440.1998
Iteration 2:  Log likelihood = -336.62162
Iteration 3:  Log likelihood = -334.64937
Iteration 4:  Log likelihood = -334.57959
Iteration 5:  Log likelihood = -334.57944
Iteration 6:  Log likelihood = -334.57944
Refining starting values:
Grid node 0:  Log likelihood = -336.03604
Fitting full model:
Iteration 0:  Log likelihood = -336.03604  (not concave)
Iteration 1:  Log likelihood = -333.14043
Iteration 2:  Log likelihood = -330.40952
Iteration 3:  Log likelihood = -329.89242
Iteration 4:  Log likelihood = -329.87847
Iteration 5:  Log likelihood = -329.87832
Iteration 6:  Log likelihood = -329.87832
Mixed-effects Weibull PH regression             Number of obs     =         76
Group variable: patient                         Number of groups  =         38
                                                Obs per group:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
Integration method: mvaghermite                 Integration pts.  =          7
                                                Wald chi2(2)      =      10.12
Log likelihood = -329.87832                     Prob > chi2       =     0.0063
```

_t	Haz. ratio	Std. err.	z	P>|z|	[95% conf.	interval]
age	1.007348	.013788	0.53	0.593	.9806828	1.034737
female	.1904727	.099992	-3.16	0.002	.0680737	.5329493
_cons	.0072901	.0072274	-4.96	0.000	.0010444	.0508881
/ln_p	.2243233	.1402795			-.0506195	.4992661
patient						
var(_cons)	.8308495	.4978425			.256735	2.688808

```
Note: Estimates are transformed only in the first equation to hazard ratios.
Note: _cons estimates baseline hazard (conditional on zero random effects).
LR test vs. Weibull model: chibar2(01) = 9.40         Prob >= chibar2 = 0.0011
```

The results are similar to those in [ST] **streg**. The likelihood-ratio test compares the random-effects model with a survival model with fixed-effects only. The results support the random-effects model.

By default, when fitting a model with the PH parameterization, `mestreg` displays exponentiated coefficients, labeled as hazard ratios. These hazard ratios should be interpreted as "conditional hazard ratios", that is, conditional on the random effects.

For example, the hazard ratio for age is 1.01. This means that according to the model, for a given patient, the hazard would increase 1% with each year of age. However, at the population level, marginal hazards corresponding to different levels of the covariates are not necessarily proportional. Example 5 in [ME] **mestreg postestimation** illustrates this point with simulated data.

The exponentiated coefficients of covariates that usually remain constant within a group do not have a natural interpretation as conditional hazard ratios. However, the magnitude of the exponentiated coefficients always gives an idea of the effect of the covariates. In this example, `female` is constant within the group. The estimated hazard ratio for `female` is 0.19, which indicates that hazard functions for females tend to be smaller than hazard functions for males. Both conditional and unconditional predictions can be obtained with `predict`. Unconditional predictions can be visualized by using `stcurve`. Unconditional effects can be tested and visualized by using `margins` and `marginsplot`. See example 1 in [ME] **mestreg postestimation** for an example using `predict`, `margins`, and `marginsplot`.

◁

▷ Example 2: Two-level random-intercept AFT model

Although the PH parameterization is more popular in the literature because the output is easier to interpret, the AFT parameterization is useful when we need to make comparisons with other models that have only an AFT parameterization. For example, we might want to compare the Weibull results from example 1 with the results from a gamma model.

Let's redisplay the results of a Weibull PH model from example 1 as coefficients:

```
. mestreg, nohr

Mixed-effects Weibull PH regression                 Number of obs     =         76
Group variable: patient                             Number of groups  =         38

                                                    Obs per group:
                                                                  min =          2
                                                                  avg =        2.0
                                                                  max =          2

Integration method: mvaghermite                     Integration pts.  =          7

                                                    Wald chi2(2)      =      10.12
Log likelihood = -329.87832                         Prob > chi2       =     0.0063

------------------------------------------------------------------------------
          _t | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         age |   .0073207   .0136874     0.53   0.593    -.0195062    .0341476
      female |  -1.658247   .5249676    -3.16   0.002    -2.687164    -.629329
       _cons |  -4.921236   .9914009    -4.96   0.000    -6.864346   -2.978126
-------------+----------------------------------------------------------------
       /ln_p |   .2243233   .1402795                     -.0506195    .4992661
-------------+----------------------------------------------------------------
patient      |
   var(_cons)|   .8308495   .4978425                       .256735    2.688808
------------------------------------------------------------------------------
LR test vs. Weibull model: chibar2(01) = 9.40         Prob >= chibar2 = 0.0011
```

We can refit the Weibull model using the AFT parameterization by specifying option time.

```
. mestreg age female || patient:, distribution(weibull) time

        Failure _d: infect
  Analysis time _t: time

Fitting fixed-effects model:

Iteration 0:  Log likelihood = -346.46486
Iteration 1:  Log likelihood = -343.29515
Iteration 2:  Log likelihood =  -335.0513
Iteration 3:  Log likelihood = -334.58308
Iteration 4:  Log likelihood = -334.57944
Iteration 5:  Log likelihood = -334.57944

Refining starting values:

Grid node 0:  Log likelihood = -335.10428

Fitting full model:

Iteration 0:  Log likelihood = -335.10428
Iteration 1:  Log likelihood = -332.13546
Iteration 2:  Log likelihood = -330.01623
Iteration 3:  Log likelihood = -329.88013
Iteration 4:  Log likelihood = -329.87832
Iteration 5:  Log likelihood = -329.87832

Mixed-effects Weibull AFT regression            Number of obs     =         76
Group variable: patient                         Number of groups  =         38

                                                Obs per group:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2

Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(2)      =      13.00
Log likelihood = -329.87832                     Prob > chi2       =     0.0015

------------------------------------------------------------------------------
          _t | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         age |  -.0058496    .010872    -0.54   0.591    -.0271585    .0154592
      female |   1.325034   .3719102     3.56   0.000      .596103    2.053964
       _cons |   3.932346   .5663757     6.94   0.000      2.82227    5.042422
-------------+----------------------------------------------------------------
       /ln_p |   .2243237   .1402794                     -.0506189    .4992663
-------------+----------------------------------------------------------------
patient      |
   var(_cons)|   .5304902   .2343675                      .2231626    1.261053
------------------------------------------------------------------------------
LR test vs. Weibull model: chibar2(01) = 9.40         Prob >= chibar2 = 0.0011
```

The estimates of coefficients and variance components are different between the two models. In fact, the coefficients have the opposite signs. This is expected because the two models have different parameterizations. The relationship between the coefficients and variances of the two parameterizations for the Weibull model is

$$\beta_{\text{PH}} = -p \times \beta_{\text{AFT}}$$

$$\text{var}_{\text{PH}} = p^2 \times \text{var}_{\text{AFT}}$$

where p denotes the ancillary parameter. It is estimated in the logarithmic metric and is displayed in the output as /ln_p.

For example, we could calculate β_{PH} for `female` as approximately $-\exp(0.22) \times 1.33 = -1.66$. If we exponentiate this to obtain the hazard ratio that was reported in example 1, we obtain the same reported result, 0.19.

For a discussion of the differences between the PH and AFT parameterizations, see, for example, Cleves, Gould, and Marchenko (2016).

Now, we can compare the results from our Weibull specification with the results from a gamma specification.

```
. mestreg age female || patient:, distribution(gamma)
        Failure _d: infect
  Analysis time _t: time
Fitting fixed-effects model:
Iteration 0:  Log likelihood = -351.17349
Iteration 1:  Log likelihood = -337.04571
Iteration 2:  Log likelihood = -335.10167
Iteration 3:  Log likelihood = -335.09115
Iteration 4:  Log likelihood = -335.09115
Refining starting values:
Grid node 0:  Log likelihood = -334.49759
Fitting full model:
Iteration 0:  Log likelihood = -334.49759
Iteration 1:  Log likelihood = -331.87827
Iteration 2:  Log likelihood = -329.64795
Iteration 3:  Log likelihood = -329.52682
Iteration 4:  Log likelihood = -329.52635
Iteration 5:  Log likelihood = -329.52634
Mixed-effects gamma AFT regression              Number of obs     =         76
Group variable: patient                         Number of groups  =         38
                                                Obs per group:
                                                              min =          2
                                                              avg =        2.0
                                                              max =          2
Integration method: mvaghermite                 Integration pts.  =          7
                                                Wald chi2(2)      =      13.23
Log likelihood = -329.52634                     Prob > chi2       =     0.0013

------------------------------------------------------------------------------
          _t | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         age |  -.0060276   .0108267    -0.56   0.578    -.0272475    .0151924
      female |   1.324745   .3685132     3.59   0.000     .6024726    2.047018
       _cons |   3.873854   .5628993     6.88   0.000     2.770592    4.977117
-------------+----------------------------------------------------------------
       /logs |  -.1835075   .1008892                     -.3812467    .0142317
-------------+----------------------------------------------------------------
patient      |
   var(_cons)|   .5071823   .2241959                       .213254    1.206232
------------------------------------------------------------------------------
LR test vs. gamma model: chibar2(01) = 11.13          Prob >= chibar2 = 0.0004
```

The coefficients and the random-effects variance are very similar for the two AFT models.

We can compare the marginal distributions or hazard functions for the two models by using stcurve; see example 2 in [ME] **mestreg postestimation**.

◁

▷ Example 3: Two-level random-slope model

In this example, we use a modified form of the dataset from Rabe-Hesketh and Skrondal (2022, sec. 15.7), previously published in Danahy et al. (1977) and analyzed by Pickles and Crouchley (1994, 1995) and Rabe-Hesketh, Skrondal, and Pickles (2004).

angina.dta includes data on 21 patients with coronary heart disease who participated in a randomized crossover trial comparing a drug to prevent angina (chest pain) with a placebo. The participants are identified by pid.

Before receiving the drug (or placebo), participants were asked to exercise on exercise bikes to the onset of angina or, if angina did not occur, to exhaustion. The exercise time, seconds, and the result of the exercise, angina—angina (angina=1) or exhaustion (angina=0)—were recorded. The drug (treat=1) or placebo (treat=0) was then taken orally, and the exercise test was repeated one, three, and five hours (variable occasion) after drug or placebo administration. Because each exercise test can have a failure (the occurrence of angina), the test is the subject. Each test is identified by tid. Failure is indicated by the variable angina. In this case, we have eight repeated measures per study participant.

Before fitting the model, we stset our data:

```
. use https://www.stata-press.com/data/r19/angina
(Angina drug data, Rabe-Hesketh and Skrondal (2021, ch. 15.7))

. stset seconds, failure(angina) id(tid)

Survival-time data settings

           ID variable: tid
         Failure event: angina!=0 & angina<.
Observed time interval: (seconds[_n-1], seconds]
     Exit on or before: failure

--------------------------------------------------------------------------
        168  total observations
          0  exclusions
--------------------------------------------------------------------------
        168  observations remaining, representing
        168  subjects
        155  failures in single-failure-per-subject data
     47,267  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =       743
```

To reiterate, we specify seconds as the time variable, angina as the failure variable, and tid as the variable identifying multiple observations per test.

Rabe-Hesketh and Skrondal (2022) apply several models to this dataset, including a lognormal model and a Cox model with random effects. We fit a Weibull model with covariates occasion and treat and interaction between occasion and treat. We include a random effect at the subject level.

```
. mestreg occasion##treat || pid:, distribution(weibull) nofvlabel
        Failure _d: angina
  Analysis time _t: seconds
       ID variable: tid
note: 1.occasion#1.treat identifies no observations in the sample.
note: 4.occasion#1.treat omitted because of collinearity.
 (output omitted)
Mixed-effects Weibull PH regression             Number of obs     =        168
Group variable: pid                             Number of groups  =         21

                                                Obs per group:
                                                              min =          8
                                                              avg =        8.0
                                                              max =          8

Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(6)      =      78.14
Log likelihood = -885.67135                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    occasion |
          2  |    .719456   .2031744    -1.17   0.244     .4136423    1.251364
          3  |    .902988   .2542476    -0.36   0.717     .5200146    1.568009
          4  |   1.264262   .3516347     0.84   0.399     .7329746    2.180648
             |
     1.treat |   .3825531    .128784    -2.85   0.004     .1977608    .7400195
             |
    occasion#|
       treat |
        1 1  |          1  (empty)
        2 1  |   .1576401   .0804767    -3.62   0.000     .0579589    .4287586
        3 1  |   .4512793   .2127706    -1.69   0.091     .1791093    1.137032
        4 1  |          1  (omitted)
             |
       _cons |   4.90e-13   9.98e-13   -13.91   0.000     9.03e-15    2.66e-11
-------------+----------------------------------------------------------------
       /ln_p |   1.640297   .0689544                      1.505149    1.775445
-------------+----------------------------------------------------------------
pid          |
   var(_cons)|   4.529641   1.544175                      2.322124    8.835725
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation to hazard ratios.
Note: _cons estimates baseline hazard (conditional on zero random effects).
LR test vs. Weibull model: chibar2(01) = 177.40       Prob >= chibar2 = 0.0000
```

Because individuals were exercising without the administration of a placebo or treatment at the first occasion (`occasion==1`), the category for interaction between `occasion==1` and `treat==1` is empty.

The estimated variance at the individual level (that is, the variance between individuals) is equal to 4.53. The likelihood-ratio test shows evidence in favor of the random-effects model versus the fixed-effects model.

The parameter p is $\exp(1.640297) = 5.16$, which is larger than 1. This means that the estimated hazard (conditional on the covariates and on the random effects) is a monotonically increasing function if we assume a Weibull distribution.

The model contains interaction terms for `occasion` and `treat`. Interpretation of interaction terms is usually less straightforward. Briefly, to interpret the exponentiated coefficients as conditional hazard ratios, we need to examine all the covariates in the interaction. The hazard function for $\texttt{pid} = j$, when we set $\texttt{occasion} = k$ and $\texttt{treat} = l$, will be

$$h(t) = h_0(t) \times \exp(\beta_{\text{occ}_k} + \beta_{\text{treat}_l} + \beta_{\text{occ}_k \times \text{treat}_l} + \texttt{\_cons} + u_j)$$

where β_{occ_k}, β_{treat_l}, and $\beta_{\text{occ}_k \times \text{treat}_l}$ are, respectively, the coefficients for the dummies for $\texttt{occasion} = k$ and $\texttt{treat} = l$ and the interaction ($\texttt{occasion} = k \times \texttt{treatment} = l$).

For example, when $\texttt{treat} = 0$, the hazard function is

$$h(t|\texttt{treat} = 0, \texttt{occasion} = k, \texttt{pid} = j) = h_0(t) \times \exp(\beta_{\text{occ}_k} + \texttt{\_cons} + u_j)$$

where β_{occ_1} is equal to 0 because $\texttt{occasion} = 1$ is the base category. This means that for a given `pid`,

$$\frac{h(t|\texttt{treat} = 0, \texttt{occ} = k, \texttt{pid} = j)}{h(t|\texttt{treat} = 0, \texttt{occ} = 1, \texttt{pid} = j)} = \exp(\beta_{\text{occ}_k})$$

Notice that this is only true within `pid`, because different participants have different u_js.

The coefficients have already been exponentiated, so we can see clearly that according to this model, when there is no treatment, the hazard for occasion 2 is smaller than the hazard for occasion 1. The increasing ratios indicate that the hazard increases with the occasion. Similar calculations could be performed for other interaction terms.

The easiest way to interpret models with interactions is by using `margins` and `marginsplot`, which allow us to compute and then visualize unconditional predictions and marginal effects. See [R] **margins** for more information.

Above we assumed a constant treatment effect for all individuals for each occasion. However, we may instead believe that the treatment effect varies also with individuals. This can be modeled by adding a random coefficient for the treatment, `i.treat`, at the individual level; we also include the `covariance(unstructured)` option to estimate a covariance term between the random intercept and the random slope for `1.treat`.

```
. mestreg occasion##treat || pid: i.treat, distribution(weibull)
> covariance(unstructured) nofvlabel

        Failure _d: angina
  Analysis time _t: seconds
       ID variable: tid
note: 1.occasion#1.treat identifies no observations in the sample.
note: 4.occasion#1.treat omitted because of collinearity.
  (output omitted)
Mixed-effects Weibull PH regression                 Number of obs     =        168
Group variable: pid                                 Number of groups  =         21

                                                    Obs per group:
                                                                  min =          8
                                                                  avg =        8.0
                                                                  max =          8

Integration method: mvaghermite                     Integration pts.  =          7

                                                    Wald chi2(6)      =      50.18
Log likelihood = -859.50038                         Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    occasion |
          2  |   .5993591   .1861745    -1.65   0.099     .3260503    1.101766
          3  |   .8643306   .2560242    -0.49   0.623      .483665    1.544597
          4  |   1.333201   .3843218     1.00   0.318     .7577392    2.345694
             |
     1.treat |   .2147751   .1280091    -2.58   0.010     .0667814    .6907365
             |
    occasion#|
       treat |
        1 1  |          1  (empty)
        2 1  |   .1594337   .0885644    -3.31   0.001     .0536714    .4736058
        3 1  |   .4632936   .2273925    -1.57   0.117     .1770402    1.212385
        4 1  |          1  (omitted)
             |
       _cons |   6.21e-17   1.75e-16   -13.20   0.000     2.44e-19    1.58e-14
-------------+----------------------------------------------------------------
       /ln_p |    1.91931   .0736166                      1.775024    2.063596
-------------+----------------------------------------------------------------
pid          |
 var(1.treat)|   4.682507   1.956897                      2.064178    10.62208
   var(_cons)|   6.939041   2.372975                      3.549852    13.56403
-------------+----------------------------------------------------------------
pid          |
 cov(1.treat,|
       _cons)|    1.73782   1.313054     1.32   0.186    -.8357182    4.311357
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation to hazard ratios.
Note: _cons estimates baseline hazard (conditional on zero random effects).
LR test vs. Weibull model: chi2(3) = 229.74               Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

We obtain somewhat different estimates of hazard ratios, but our inferential conclusions remain the same. We now observe two variances in the output, the variance for the intercept at the individual level and the variance for the coefficient for treatment at the individual level. The variance for the intercept is smaller because some of the variability is now explained by varying coefficients for treatment. The covariance is positive, meaning that the random slope tends to be larger for individuals who have a larger random intercept. See example 4 in [ME] **mestreg postestimation** for an application of `predict` that presents a graphical analysis of this relationship.

◁

Three-level models

▷ Example 4: Three-level random-slope model

Blossfeld, Golsch, and Rohwer (2007) analyze a dataset based on the German Life History Study of Mayer and Brückner (1989), collected in the years 1981–1983. (This dataset is also available in Blossfeld, Rohwer, and Schneider (2019), a second edition of the 2007 reference.) The `jobhistory` dataset contains a modified version of Blossfeld, Golsch, and Rohwer's anonymization of a random sample of 201 respondents from the original data. Each of the 600 observations in the dataset corresponds to a job episode. Variable `id` contains identification of the individual, `tstart` contains the starting point of the job (in months from the beginning of the century), `tend` is the end of the job episode, and `failure` indicates whether the date in `tend` corresponds to the actual end of the employment in a certain job or whether it is a censored observation.

We first `stset` the data. As explained in Cleves (1999) and Therneau and Grambsch (2000), when analyzing multiple-failure data, we can consider two main approaches. One approach is to define the study time from the first time that an individual starts being at risk. The second approach is to define the study time from the last failure. We will take the second approach, which means that we treat each job episode as the subject.

Therefore, the origin is defined as the start of each job episode, and the study time will be the time from the start of each episode until the jobs end or the episode is censored.

```
. use https://www.stata-press.com/data/r19/jobhistory
(Job history data, Event History Analysis with Stata, Blossfeld et al. 2007)

. stset tend, origin(tstart) failure(failure)

Survival-time data settings

         Failure event: failure!=0 & failure<.
Observed time interval: (origin, tend]
     Exit on or before: failure
     Time for analysis: (time-origin)
                Origin: time tstart
──────────────────────────────────────────────────────────────────────────
        600  total observations
          0  exclusions
──────────────────────────────────────────────────────────────────────────
        600  observations remaining, representing
        458  failures in single-record/single-failure data
     40,782  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =       428
```

We want to fit a Weibull model using the education level, the number of previous jobs, the prestige of the current job, and gender as explanatory variables. `education` records the highest education level before entering the labor market, `njobs` contains the number of previous jobs for each individual, and `prestige` is an index for the prestige of the current job. The `birthyear` variable indicates the year of birth. `female` is 1 for women, 0 for men. To account for individual heterogeneity, we include a random effect at the individual level.

```
. mestreg education njobs prestige i.female || id:, distribution(weibull)

        Failure _d: failure
  Analysis time _t: (tend-origin)
            Origin: time tstart

Fitting fixed-effects model:

Iteration 0:  Log likelihood = -5736904.5
Iteration 1:  Log likelihood = -2664.7487
Iteration 2:  Log likelihood = -2484.7829
Iteration 3:  Log likelihood = -2477.4358
Iteration 4:  Log likelihood = -2477.3338
Iteration 5:  Log likelihood = -2477.3337

Refining starting values:

Grid node 0:  Log likelihood = -2491.2191

Fitting full model:

Iteration 0:  Log likelihood = -2491.2191  (not concave)
Iteration 1:  Log likelihood = -2468.3995
Iteration 2:  Log likelihood = -2450.0938
Iteration 3:  Log likelihood = -2443.0739
Iteration 4:  Log likelihood =  -2442.875
Iteration 5:  Log likelihood = -2442.8747
Iteration 6:  Log likelihood = -2442.8746

Mixed-effects Weibull PH regression              Number of obs     =        600
Group variable: id                               Number of groups  =        201

                                                 Obs per group:
                                                               min =          1
                                                               avg =        3.0
                                                               max =          9

Integration method: mvaghermite                  Integration pts.  =          7

                                                 Wald chi2(4)      =      87.38
Log likelihood = -2442.8746                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
   education |    1.11897   .0463468     2.71   0.007     1.031722    1.213597
       njobs |   .7071195   .0357624    -6.85   0.000     .6403884    .7808043
    prestige |   .9677567   .0069576    -4.56   0.000     .9542157      .98149
    1.female |    1.75651   .3185526     3.11   0.002     1.231063    2.506228
       _cons |   .0053352   .0029015    -9.62   0.000     .0018376    .0154904
-------------+----------------------------------------------------------------
       /ln_p |   .1695545   .0453649                      .0806409    .2584681
-------------+----------------------------------------------------------------
id           |
   var(_cons)|   1.016459   .2149037                       .671623    1.538347
------------------------------------------------------------------------------
Note: Estimates are transformed only in the first equation to hazard ratios.
Note: _cons estimates baseline hazard (conditional on zero random effects).
LR test vs. Weibull model: chibar2(01) = 68.92        Prob >= chibar2 = 0.0000
```

The estimated variance of the random intercept is equal to 1.02.

According to this model, an increase in the number of previous jobs is negatively associated with job mobility; the same is true for an increase in the prestige of the current job. By contrast, an increase in the years of education is positively associated with job mobility. Also, women seem to be more mobile than men.

We now store our estimates for later use:

```
. estimates store randint
```

The dataset has only two natural levels. However, for illustration purposes, let's consider the following situation. Assume that we want to account for unobserved variables associated with the date of birth, such as life experience, level of familiarity with new technologies, and family situation. We therefore add a random effect for the year of birth. Now, individuals will be nested within birth years.

```
. mestreg education njobs prestige i.female || birthyear: || id:,
> distribution(weibull)

        Failure _d: failure
  Analysis time _t: (tend-origin)
            Origin: time tstart
```

(output omitted)

```
Mixed-effects Weibull PH regression             Number of obs     =        600

        Grouping information
```

Group variable	No. of groups	Observations per group Minimum	Average	Maximum
birthyear	12	3	50.0	99
id	201	1	3.0	9

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(4)      =      83.20
Log likelihood = -2439.9066                     Prob > chi2       =     0.0000
```

_t	Haz. ratio	Std. err.	z	P>|z|	[95% conf.	interval]
education	1.120373	.045203	2.82	0.005	1.035189	1.212566
njobs	.7181197	.0372039	-6.39	0.000	.6487813	.7948686
prestige	.966567	.0069189	-4.75	0.000	.9531009	.9802234
1.female	1.734236	.3022479	3.16	0.002	1.232419	2.440384
_cons	.0059091	.0031758	-9.55	0.000	.0020609	.0169429
/ln_p	.1685641	.0454824			.0794203	.257708
birthyear						
var(_cons)	.0950371	.0741445			.0205976	.4385006
birthyear>id						
var(_cons)	.8728384	.2020938			.5544339	1.374099

```
Note: Estimates are transformed only in the first equation to hazard ratios.
Note: _cons estimates baseline hazard (conditional on zero random effects).
LR test vs. Weibull model: chi2(2) = 74.85                Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

The results for the fixed part of the model are similar to the ones in the previous model.

Now, we have two estimated variances—one estimate for the random intercept at the individual level and one estimate for the random intercept at the birth-year level.

The variance component for the individual level is smaller for this model, and it looks as if the first model might have been trying to explain a variance component at the birth-year level by incorporating it into the individual-level variance. We can perform a likelihood-ratio test to compare the stored model `randint` with the current model:

```
. lrtest randint .
Likelihood-ratio test
Assumption: randint nested within .
 LR chi2(1) =   5.94
Prob > chi2 = 0.0148
Note: The reported degrees of freedom assumes the null hypothesis is not on
      the boundary of the parameter space. If this is not true, then the
      reported test is conservative.
```

The test is conservative because we are testing on the boundary of the parameter space; see *Distribution theory for likelihood-ratio test* in [ME] **me** for details. Provided that we are testing only one variance component, we can adjust the p-value accordingly by dividing the reported value by two, which results in an adjusted p-value equal to 0.0074.

The test is significant at the 0.05 level. It supports the three-level model with the additional variance component at the birth-year level.

◁

Stored results

`mestreg` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(k)`	number of parameters
`e(k_eq)`	number of equations in `e(b)`
`e(k_eq_model)`	number of equations in overall model test
`e(k_dv)`	number of dependent variables
`e(k_f)`	number of fixed-effects parameters
`e(k_r)`	number of random-effects parameters
`e(k_rs)`	number of variances
`e(k_rc)`	number of covariances
`e(df_m)`	model degrees of freedom
`e(ll)`	log likelihood
`e(chi2)`	χ^2
`e(p)`	p-value for model test
`e(ll_c)`	log likelihood, comparison model
`e(chi2_c)`	χ^2, comparison test
`e(df_c)`	degrees of freedom, comparison test
`e(p_c)`	p-value for comparison test
`e(N_clust)`	number of clusters
`e(rank)`	rank of `e(V)`
`e(ic)`	number of iterations
`e(rc)`	return code
`e(converged)`	1 if converged, 0 otherwise

Macros

`e(cmd)`	`gsem`
`e(cmd2)`	`mestreg`
`e(cmdline)`	command as typed
`e(depvar)`	name of dependent variable
`e(wtype)`	weight type

`e(wexp)`	weight expression (first-level weights)
`e(fweight`*k*`)`	`fweight` variable for *k*th highest level, if specified
`e(iweight`*k*`)`	`iweight` variable for *k*th highest level, if specified
`e(pweight`*k*`)`	`pweight` variable for *k*th highest level, if specified
`e(covariates)`	list of covariates
`e(ivars)`	grouping variables
`e(model)`	model name
`e(title)`	title in estimation output
`e(distribution)`	distribution
`e(clustvar)`	name of cluster variable
`e(offset)`	offset
`e(intmethod)`	integration method
`e(n_quad)`	number of integration points
`e(chi2type)`	`Wald`; type of model χ^2
`e(vce)`	*vcetype* specified in `vce()`
`e(vcetype)`	title used to label Std. err.
`e(frm2)`	`hazard` or `time`
`e(opt)`	type of optimization
`e(which)`	`max` or `min`; whether optimizer is to perform maximization or minimization
`e(ml_method)`	type of `ml` method
`e(user)`	name of likelihood-evaluator program
`e(technique)`	maximization technique
`e(datasignature)`	the checksum
`e(datasignaturevars)`	variables used in calculation of checksum
`e(properties)`	`b V`
`e(estat_cmd)`	program used to implement `estat`
`e(predict)`	program used to implement `predict`
`e(marginsnotok)`	predictions disallowed by `margins`
`e(marginswtype)`	weight type for `margins`
`e(marginswexp)`	weight expression for `margins`
`e(asbalanced)`	factor variables `fvset` as `asbalanced`
`e(asobserved)`	factor variables `fvset` as `asobserved`
Matrices	
`e(b)`	coefficient vector
`e(Cns)`	constraints matrix
`e(ilog)`	iteration log (up to 20 iterations)
`e(gradient)`	gradient vector
`e(N_g)`	group counts
`e(g_min)`	group-size minimums
`e(g_avg)`	group-size averages
`e(g_max)`	group-size maximums
`e(V)`	variance–covariance matrix of the estimators
`e(V_modelbased)`	model-based variance
Functions	
`e(sample)`	marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices	
`r(table)`	matrix containing the coefficients with their standard errors, test statistics, *p*-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Methods and formulas are presented under the following headings:

Survival models

Survival models have a trivariate response (t_0, t, d):

t_0 is the starting time under observation $t_0 \geq 0$;

t is the ending time under observation $t \geq t_0$; and

d is an indicator for failure $d \in \{0, 1\}$.

The survival function for a given family is the complement of the cumulative distribution function, $S(t) = 1 - F(t)$. The unconditional density for a failure at time t is given by

$$g(t) = \frac{\partial F(t)}{\partial t} = -\frac{\partial S(t)}{\partial t}$$

Some distributions contain ancillary parameters that are not denoted here.

The conditional density for a failure at time t is

$$g(t|t \geq t_0, d = 1) = g(t)/S(t_0)$$

and the conditional probability of survival without failure up to time t is

$$P(T \geq t|t \geq t_0, d = 0) = S(t)/S(t_0)$$

The conditional likelihood is given by

$$L(t, t_0, d) = \left\{ \frac{g(t)}{S(t_0)} \right\}^d \left\{ \frac{S(t)}{S(t_0)} \right\}^{1-d}$$

See *Survival distributions* in [SEM] **Methods and formulas for gsem** for the specific density function corresponding to each distribution.

Given a set of cluster-level random effects $\mathbf{u}_j$ for $j = 1, \ldots, M$, the conditional distribution of $\mathbf{t}_j = (t_{j1}, \ldots, t_{jn_j})'$ on $\boldsymbol{\eta}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j = (\mathbf{x}_{j1}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j, \ldots, \mathbf{x}_{jn_j}\boldsymbol{\beta} + \mathbf{z}_{jn_j}\mathbf{u}_j)$ for cluster j is

$$f(\mathbf{t}_j|\boldsymbol{\eta}_j) = \prod_{i=1}^{n_j} f(t_{ji}|\eta_{ji})$$

where $f(t_{ji}|\eta_{ji})$ is the contribution to the likelihood from observation ji; that is,

$$f(t_{ji}|\eta_{ji}) = \left\{ \frac{g(t_{ji}|\mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j)}{S(t_{0ji}|\mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j)} \right\}^{d_{ji}} \left\{ \frac{S(t_{ji}|\mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j)}{S(t_{0ji}|\mathbf{x}_{ji}\boldsymbol{\beta} + \mathbf{z}_{ji}\mathbf{u}_j)} \right\}^{1-d_{ji}} \tag{1}$$

where $g(t|\eta)$ and $S(t|\eta)$ are, respectively, the density and the survivor function conditional on the linear prediction η.

As mentioned in *Introduction* under *Remarks and examples*, `mestreg` does not allow delayed entry or gaps. Therefore, the first observation for a given subject will have a value of $t_0 = 0$, and subsequent spells for the subject must start at the end of the previous spell. That is, if observations ji and $j, i+1$ belong to the same subject, then $t_{0j,i+1} = t_{ji}$.

Because the prior distribution of $\mathbf{u}_j$ is multivariate normal with mean $\mathbf{0}$ and $q \times q$ variance matrix $\mathbf{\Sigma}$, the likelihood contribution for the jth cluster is obtained by integrating $\mathbf{u}_j$ out of the joint density $f(\mathbf{t}_j, \mathbf{u}_j)$,

$$\mathcal{L}_j(\boldsymbol{\beta}, \mathbf{\Sigma}) = (2\pi)^{-q/2} \left|\mathbf{\Sigma}\right|^{-1/2} \int f(\mathbf{t}_j | \mathbf{X}_j \boldsymbol{\beta} + \mathbf{Z}_j \mathbf{u}_j) \exp\left(-\mathbf{u}_j' \mathbf{\Sigma}^{-1} \mathbf{u}_j / 2\right) d\mathbf{u}_j \tag{2}$$

The integration in (2) has no closed form and thus must be approximated; see *Methods and formulas* in [ME] **meglm** for details.

Survey data

In the presence of sampling weights, following Rabe-Hesketh and Skrondal (2006), the weighted log pseudolikelihood for a two-level model is given as

$$\mathcal{L}(\boldsymbol{\beta}, \mathbf{\Sigma}) = \sum_{j=1}^{M} w_j \log \int_{-\infty}^{\infty} \exp\left\{ \sum_{i=1}^{n_j} w_{i|j} \log f(t_{ji} | \eta_{ji}) \right\} \phi(\mathbf{v}_{j1})\, d\mathbf{v}_{j1}$$

where w_j is the inverse of the probability of selection for the jth cluster; $w_{i|j}$ is the inverse of the conditional probability of selection of individual i, given the selection of cluster j; $f(t_{ji}|\eta_{ji})$ is as in (1); and η_{ji}, $\phi(\cdot)$, $\mathbf{v}_{j1}$ are defined as in *Methods and formulas* in [ME] **meglm**.

Weighted estimation is achieved through the direct application of w_j and $w_{i|j}$ into the likelihood calculations as detailed above to reflect replicated clusters for w_j and replicated observations within clusters for $w_{i|j}$. Because this estimation is based on replicated clusters and observations, frequency weights are handled similarly.

References

Andrews, M. J., T. Schank, and R. Upward. 2006. Practical fixed-effects estimation methods for the three-way error-components model. *Stata Journal* 6: 461–481.

Blossfeld, H.-P., K. Golsch, and G. Rohwer. 2007. *Event History Analysis with Stata*. Mahwah, NJ: Erlbaum.

Blossfeld, H.-P., G. Rohwer, and T. Schneider. 2019. *Event History Analysis with Stata*. 2nd ed. New York: Routledge.

Cleves, M. A. 1999. FAQ: How do I analyze multiple failure-time data using Stata? https://www.stata.com/support/faqs/statistics/multiple-failure-time-data.

Cleves, M. A., W. W. Gould, and Y. V. Marchenko. 2016. *An Introduction to Survival Analysis Using Stata*. Rev. 3rd ed. College Station, TX: Stata Press.

Crowther, M. J. 2019. Multilevel mixed-effects parametric survival analysis: Estimation, simulation, and application. *Stata Journal* 19: 931–949.

Danahy, D. T., D. T. Burwell, W. S. Aronow, and R. Prakash. 1977. Sustained hemodynamic and antianginal effect of high dose oral isosorbide dinitrate. *Circulation* 55: 381–387. https://doi.org/10.1161/01.cir.55.2.381.

Joe, H. 2008. Accuracy of Laplace approximation for discrete response mixed models. *Computational Statistics and Data Analysis* 52: 5066–5074. https://doi.org/10.1016/j.csda.2008.05.002.

Laird, N. M., and J. H. Ware. 1982. Random-effects models for longitudinal data. *Biometrics* 38: 963–974. https://doi.org/10.2307/2529876.

Leyland, A. H., and H. Goldstein, eds. 2001. *Multilevel Modelling of Health Statistics*. New York: Wiley.

Lin, X., and N. E. Breslow. 1996. Bias correction in generalized linear mixed models with multiple components of dispersion. *Journal of the American Statistical Association* 91: 1007–1016. https://doi.org/10.2307/2291720.

Marchenko, Y. V. 2006. Estimating variance components in Stata. *Stata Journal* 6: 1–21.

Mayer, K. U., and E. Brückner. 1989. Lebensverläufe und Wohlfahrtsentwicklung: Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929–1931, 1939–1941, 1949–1951. Materialien aus der Bildungsforschung 35, Max-Planck-Institut für Bildungsforschung, Berlin.

McCulloch, C. E., S. R. Searle, and J. M. Neuhaus. 2008. *Generalized, Linear, and Mixed Models*. 2nd ed. Hoboken, NJ: Wiley.

McLachlan, G. J., and K. E. Basford. 1988. *Mixture Models: Inference and Applications to Clustering*. New York: Dekker.

Pickles, A., and R. Crouchley. 1994. Generalizations and applications of frailty models for survival and event data. *Statistical Methods in Medical Research* 3: 263–278. https://doi.org/10.1177/096228029400300305.

———. 1995. A comparison of frailty models for multivariate survival data. *Statistics in Medicine* 14: 1447–1461. https://doi.org/10.1002/sim.4780141305.

Rabe-Hesketh, S., and A. Skrondal. 2006. Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society*, A ser., 169: 805–827. https://doi.org/10.1111/j.1467-985X.2006.00426.x.

———. 2022. *Multilevel and Longitudinal Modeling Using Stata*. 4th ed. College Station, TX: Stata Press.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles. 2004. Generalized multilevel structural equation modeling. *Psychometrika* 69: 167–190. https://doi.org/10.1007/BF02295939.

———. 2005. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics* 128: 301–323. https://doi.org/10.1016/j.jeconom.2004.08.017.

Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage.

Self, S. G., and K.-Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* 82: 605–610. https://doi.org/10.2307/2289471.

Skrondal, A., and S. Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: Chapman and Hall/CRC.

Therneau, T. M., and P. M. Grambsch. 2000. *Modeling Survival Data: Extending the Cox Model*. New York: Springer.

Also see

[ME] **mestreg postestimation** — Postestimation tools for mestreg

[ME] **me** — Introduction to multilevel mixed-effects models

[BAYES] **bayes: mestreg** — Bayesian multilevel parametric survival models

[ST] **streg** — Parametric survival models

[ST] **Glossary**

[SVY] **svy estimation** — Estimation commands for survey data

[XT] **xtstreg** — Random-effects parametric survival models

[U] 20 Estimation and postestimation commands

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