fmm: glm — Finite mixtures of generalized linear regression models

Description

`fmm: glm` fits mixtures of generalized linear regression models; see [FMM] **fmm** and [R] **glm** for details.

Quick start

Mixture of two normal distributions of `y`

```
fmm 2: glm y, family(gaussian) link(identity)
```

Mixture of two gamma distributions of `y`

```
fmm 2: glm y, family(gamma)
```

Mixture of two gamma regression models of `y` on `x1` and `x2`

```
fmm 2: glm y x1 x2, family(gamma)
```

Same as above, but with class probabilities depending on `z1` and `z2`

```
fmm 2, lcprob(z1 z2): glm y x1 x2, family(gamma)
```

With robust standard errors

```
fmm 2, vce(robust): glm y x1 x2, family(gamma)
```

Constrain coefficients on `x1` and `x2` to be equal across classes

```
fmm 2, lcinvariant(coef): glm y x1 x2
```

Menu

Statistics > FMM (finite mixture models) > Generalized linear model (GLM)

Syntax

Basic syntax

```
fmm # : glm depvar [indepvars] [ , options]
```

Full syntax

```
fmm # [if] [in] [weight] [ , fmmopts] : glm depvar [indepvars] [ , options]
```

where # specifies the number of class models.

options	Description
Model	
`family(`*familyname*`)`	distribution of *depvar*; default is `family(gaussian)`
`link(`*linkname*`)`	link function; default varies per family
`noconstant`	suppress the constant term
`exposure(`*varname*$_e$`)`	include $\ln($*varname*$_e)$ in model with coefficient constrained to 1
`offset(`*varname*$_o$`)`	include *varname*$_o$ in model with coefficient constrained to 1
`asis`	retain perfect predictor variables

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

depvar and *indepvars* may contain time-series operators; see **[U] 11.4.4 Time-series varlists**.

For a detailed description of *options*, see *Options* in [R] **glm**.

familyname	Description
`gaussian`	Gaussian (normal); the default
`bernoulli`	Bernoulli
`beta`	beta
`binomial` [# | *varname*]	binomial; default number of binomial trials is 1
`poisson`	Poisson
`nbinomial` [`mean` | `constant`]	negative binomial; default dispersion is `mean`
`exponential`	exponential
`gamma`	gamma
`lognormal`	lognormal
`loglogistic`	loglogistic
`weibull`	Weibull

`bernoulli`, `beta`, `exponential`, `lognormal`, `loglogstic`, and `weibull` are extensions available with `fmm: glm` that are not available with `glm`.

linkname	Description
`identity`	identity
`log`	log
`logit`	logit
`probit`	probit
`cloglog`	complementary log–log

fmmopts	Description
Model	
`lcinvariant(`*pclassname*`)`	specify parameters that are equal across classes; default is `lcinvariant(none)`
`lcprob(`*varlist*`)`	specify independent variables for class probabilities
`lclabel(`*name*`)`	name of the categorical latent variable; default is `lclabel(Class)`
`lcbase(`#`)`	base latent class
`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)`
`nocnsreport`	do not display constraints
`noheader`	do not display header above parameter table
`nodvheader`	do not display dependent variables information in the header
`notable`	do not display parameter table
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process
`startvalues(`*svmethod*`)`	method for obtaining starting values; default is `startvalues(factor)`
`emopts(`*maxopts*`)`	control EM algorithm for improved starting values
`noestimate`	do not fit the model; show starting values instead
`collinear`	keep collinear variables
`coeflegend`	display legend instead of statistics

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

`by`, `collect`, `statsby`, and `svy` are allowed; see **[U] 11.1.10 Prefix commands**.

`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**.

`collinear` and `coeflegend` do not appear in the dialog box.

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

For a detailed description of *fmmopts*, see *Options* in [FMM] **fmm**.

pclassname	Description
`cons`	intercepts and cutpoints
`coef`	fixed coefficients
`errvar`	covariances of errors
`scale`	scaling parameters
`all`	all the above
`none`	none of the above; the default

Remarks and examples

For a general introduction to finite mixture models, see [FMM] **fmm intro**. For general information about generalized linear regression, see [R] **glm**. For examples using `fmm`, see examples in *Contents*.

If you specify both `family()` and `link()`, not all combinations make sense. You may choose from the following combinations:

	identity	log	logit	probit	cloglog
Gaussian	D	x			
Bernoulli			D	x	x
beta			D	x	x
binomial			D	x	x
Poisson		D			
negative binomial		D			
exponential		D			
gamma		D			
lognormal		D			
loglogistic		D			
Weibull		D			

D denotes the default.

Stored results

See *Stored results* in [FMM] **fmm**.

Methods and formulas

See *Methods and formulas* in [FMM] **fmm**.

Also see

[FMM] **fmm** — Finite mixture models using the fmm prefix

[FMM] **fmm intro** — Introduction to finite mixture models

[FMM] **fmm postestimation** — Postestimation tools for fmm

[FMM] **Glossary**

[R] **glm** — Generalized linear models

[SEM] **gsem** — Generalized structural equation model estimation command

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