xtpoisson — Fixed-effects, random-effects, and population-averaged Poisson models

Description

xtpoisson fits random-effects, conditional fixed-effects, and population-averaged Poisson models. These models are typically used for a nonnegative count dependent variable.

Quick start

Random-effects Poisson regression of \( y \) on \( x \) and indicators for levels of categorical variable \( a \) using \texttt{xtset} \; \texttt{data}
\begin{verbatim}
  xtpoisson \( y \) \( x \) i.a
\end{verbatim}
Conditioned fixed-effects model with exposure variable \( evar \)
\begin{verbatim}
  xtpoisson \( y \) \( x \) i.a, fe exposure(evar)
\end{verbatim}
Population-averaged model with robust standard errors
\begin{verbatim}
  xtpoisson \( y \) \( x \) i.a, pa vce(robust)
\end{verbatim}
As above, but report incidence-rate ratios
\begin{verbatim}
  xtpoisson \( y \) \( x \) i.a, pa vce(robust) irr
\end{verbatim}

Menu

Statistics \; \textgreater \; Longitudinal/panel data \; \textgreater \; Count outcomes \; \textgreater \; Poisson regression (FE, RE, PA)
## Syntax

### Random-effects (RE) model

```
xtpoisson depvar [ indepvars ] [ if ] [ in ] [ weight ] [ , re RE_options ]
```

### Conditional fixed-effects (FE) model

```
xtpoisson depvar [ indepvars ] [ if ] [ in ] [ weight ] , fe [ FE_options ]
```

### Population-averaged (PA) model

```
xtpoisson depvar [ indepvars ] [ if ] [ in ] [ weight ] , pa [ PA_options ]
```

### RE_options Description

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>noconstant</td>
<td>suppress constant term</td>
</tr>
<tr>
<td>re</td>
<td>use random-effects estimator; the default</td>
</tr>
<tr>
<td>exposure(varname)</td>
<td>include ln(varname) in model with coefficient constrained to 1</td>
</tr>
<tr>
<td>offset(varname)</td>
<td>include varname in model with coefficient constrained to 1</td>
</tr>
<tr>
<td>normal</td>
<td>use a normal distribution for random effects instead of gamma</td>
</tr>
<tr>
<td>constraints(constraints)</td>
<td>apply specified linear constraints</td>
</tr>
</tbody>
</table>

### SE/Robust

```
vce(vcetype) vcetype may be oim, robust, cluster clustvar, bootstrap, or jackknife
```

### Reporting

```
level(#) set confidence level; default is level(95)
irr report incidence-rate ratios
llrmodel perform the likelihood-ratio model test instead of the default Wald test
noconsreport do not display constraints
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; intmethod may be mvaghermite (the default) or ghermite
intpoints(#) use # quadrature points; default is intpoints(12)
```

### Maximization

```
maximize_options control the maximization process; seldom used
collinear keep collinear variables
coefflegend display legend instead of statistics
```
### Model
- **fe**
  - Description: use fixed-effects estimator
- **exposure(varname)**
  - Description: include \(\ln(varname)\) in model with coefficient constrained to 1
- **offset(varname)**
  - Description: include \(varname\) in model with coefficient constrained to 1
- **constraints(constraints)**
  - Description: apply specified linear constraints

### SE/Robust
- **vce(vcetype)**
  - Description: \(vcetype\) may be oim, robust, bootstrap, or jackknife

### Reporting
- **level(#)***
  - Description: set confidence level; default is level(95)
- **irr**
  - Description: report incidence-rate ratios
- **nocnsreport**
  - Description: do not display constraints
- **display_options**
  - Description: 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**
  - Description: control the maximization process; seldom used
- **collinear**
  - Description: keep collinear variables
- **coeflegend**
  - Description: display legend instead of statistics
## PA_options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>noconstant</td>
<td>suppress constant term</td>
</tr>
<tr>
<td>pa</td>
<td>use population-averaged estimator</td>
</tr>
<tr>
<td>exposure(varname)</td>
<td>include ln(varname) in model with coefficient constrained to 1</td>
</tr>
<tr>
<td>offset(varname)</td>
<td>include varname in model with coefficient constrained to 1</td>
</tr>
</tbody>
</table>

## Correlation

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>corr</em>(correlation)</td>
<td>within-panel correlation structure</td>
</tr>
<tr>
<td><em>force</em></td>
<td>estimate even if observations unequally spaced in time</td>
</tr>
</tbody>
</table>

## SE/Robust

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>vce</em>(vcetype)</td>
<td>vcetype may be conventional, robust, bootstrap, or jackknife</td>
</tr>
<tr>
<td><em>nmp</em></td>
<td>use divisor $N - P$ instead of the default $N$</td>
</tr>
<tr>
<td><em>scale</em>(parm)</td>
<td>overrides the default scale parameter; parm may be $x^2$, dev, phi, or #</td>
</tr>
</tbody>
</table>

## Reporting

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>level</em>(#)</td>
<td>set confidence level; default is level(95)</td>
</tr>
<tr>
<td><em>irr</em></td>
<td>report incidence-rate ratios</td>
</tr>
<tr>
<td><em>display_options</em></td>
<td>control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling</td>
</tr>
</tbody>
</table>

## Optimization

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>optimize_options</em></td>
<td>control the optimization process; seldom used</td>
</tr>
<tr>
<td><em>coeflegend</em></td>
<td>display legend instead of statistics</td>
</tr>
</tbody>
</table>

## correlation

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>exchangeable</td>
<td>exchangeable</td>
</tr>
<tr>
<td>independent</td>
<td>independent</td>
</tr>
<tr>
<td>unstructured</td>
<td>unstructured</td>
</tr>
<tr>
<td>fixed matname</td>
<td>user-specified</td>
</tr>
<tr>
<td>ar #</td>
<td>autoregressive of order #</td>
</tr>
<tr>
<td>stationary #</td>
<td>stationary of order #</td>
</tr>
<tr>
<td>nonstationary #</td>
<td>nonstationary of order #</td>
</tr>
</tbody>
</table>

A panel variable must be specified. For xtpoisson, pa, correlation structures other than exchangeable and independent require that a time variable also be specified. Use xtset; see [XT] xtset.

depvars 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.
bayes, by, collect, mi estimate, and statsby are allowed; see [U] 11.1.10 Prefix commands. For more details, see [BAYES] bayes: xtpoisson. fp is allowed for the random-effects and fixed-effects models.
vce(bootstrap) and vce(jackknife) are not allowed with the mi estimate prefix; see [MI] mi estimate.
iweights, fweights, and pweights are allowed for the population-averaged model, and iweights are allowed for the random-effects and fixed-effects models; see [U] 11.1.6 weight. Weights must be constant within panel.
collinear and coeflegend do not appear in the dialog box.

See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.
Options for RE model

Model

*noconstant*; see [R] Estimation options.

*re*, the default, requests the random-effects estimator.

*exposure(varname)*, *offset(varname)*; see [R] Estimation options.

*normal* specifies that the random effects follow a normal distribution instead of a gamma distribution.

*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*), that are robust to some kinds of misspecification (*robust*), that allow for intragroup correlation (*cluster clustvar*), and that use bootstrap or jackknife methods (*bootstrap, jackknife*); see [XT] vce_options.

Specifying *vce(robust)* is equivalent to specifying *vce(cluster panelvar)*; see xtpoisson, re and the robust VCE estimator in Methods and formulas.

Reporting

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

*i*rr reports exponentiated coefficients $e^b$ rather than coefficients $b$. For the Poisson model, exponentiated coefficients are interpreted as incidence-rate ratios.

*lrmodel*, *nocnsrep*rt; see [R] Estimation options.

*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)*, *intpoints(#)*width; see [R] Estimation options. *normal* must also be specified.

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. These options are seldom used.

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

*collinear*, *coeflegend*; see [R] Estimation options.
Options for FE model

Model

*fe* requests the fixed-effects estimator.

```plaintext
exposure(varname), offset(varname), 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*), that are robust to some kinds of misspecification (*robust*), and that use bootstrap or jackknife methods (*bootstrap, jackknife*); see [XT] *vce_options.*

*vce(robust)* invokes a cluster–robust estimate of the VCE in which the ID variable specifies the clusters.

Reporting

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

*iir* reports exponentiated coefficients $e^b$ rather than coefficients $b$. For the Poisson model, exponentiated coefficients are interpreted as incidence-rate ratios.

*nocnsreport*; see [R] Estimation options.

*display_options*: *noci*, *nopvalues*, *noomitted*, *vsquish*, *noemptycells*, *baselevels*,

`allbaselevels, nofvlabel, fwrap(#), fwrapon(style), cformat(%,fmt), pformat(%,fmt), sformat(%,fmt), and nolstretch; see [R] Estimation options.*

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. These options are seldom used.

The following options are available with *xtpoisson* but are not shown in the dialog box:

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

Options for PA model

Model

*noconstant*; see [R] Estimation options.

*pa* requests the population-averaged estimator.

```plaintext
exposure(varname), offset(varname); see [R] Estimation options.
```

Correlation

*corr(correlation)* specifies the within-panel correlation structure; the default corresponds to the equal-correlation model, *corr(exchangeable).*

When you specify a correlation structure that requires a lag, you indicate the lag after the structure’s name with or without a blank; for example, *corr(ar 1)* or *corr(ar1).*
If you specify the fixed correlation structure, you specify the name of the matrix containing the assumed correlations following the word fixed, for example, corr(fixed myr).

force specifies that estimation be forced even though the time variable is not equally spaced. This is relevant only for correlation structures that require knowledge of the time variable. These correlation structures require that observations be equally spaced so that calculations based on lags correspond to a constant time change. If you specify a time variable indicating that observations are not equally spaced, the (time dependent) model will not be fit. If you also specify force, the model will be fit, and it will be assumed that the lags based on the data ordered by the time variable are appropriate.

\[ \text{vce(vcetype)} \] specifies the type of standard error reported, which includes types that are derived from asymptotic theory (conventional), that are robust to some kinds of misspecification (robust), and that use bootstrap or jackknife methods (bootstrap, jackknife); see [XT] vce_options.

vce(conventional), the default, uses the conventionally derived variance estimator for generalized least-squares regression.

\[ \text{nmp, scale(x2 | dev | phi | #)} \]; see [XT] vce_options.

\[ \text{level(#)} \]; see [R] Estimation options.

irr reports exponentiated coefficients \( e^b \) rather than coefficients \( b \). For the Poisson model, exponentiated coefficients are interpreted as incidence-rate ratios.

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

\[ \text{optimize_options} \] control the iterative optimization process. These options are seldom used.

iterate(#) specifies the maximum number of iterations. When the number of iterations equals #, the optimization stops and presents the current results, even if convergence has not been reached. The default is iterate(100).

tolerance(#) specifies the tolerance for the coefficient vector. When the relative change in the coefficient vector from one iteration to the next is less than or equal to #, the optimization process is stopped. tolerance(1e-6) is the default.

log and nolog specify whether to display the iteration log. The iteration log is displayed by default unless you used set iterlog off to suppress it; see set iterlog in [R] set iter.

trace specifies that the current estimates be printed at each iteration.

The following option is available with xtpoisson but is not shown in the dialog box: coeflegend; see [R] Estimation options.
Remarks and examples

_xtpoisson_ fits random-effects, conditional fixed-effects, and population-averaged Poisson models. Whenever we refer to a fixed-effects model, we mean the conditional fixed-effects model. These models are typically used for a nonnegative count dependent variable but may be used for any dependent variable in natural logs. For more information about the assumptions of the Poisson model, see [R] _poisson_.

By default, the population-averaged model is an equal-correlation model; _xtpoisson_, _pa_ assumes 

**corr(exchangeable).** Thus, _xtpoisson_ is a convenience command for fitting the population-averaged model using _xtgee_; see [XT] _xtgee_. Typing

```
  . xtpoisson ..., ... pa exposure(time)
```

is equivalent to typing

```
  . xtgee ..., ... family(poisson) link(log) corr(exchangeable) exposure(time)
```

Also see [XT] _xtgee_ for information about _xtpoisson_.

By default or when _re_ is specified, _xtpoisson_ fits via maximum likelihood the random-effects model

\[
\Pr(Y_{it} = y_{it}|x_{it}) = F(y_{it}, x_{it} \beta + \nu_i)
\]

for \(i = 1, \ldots, n\) panels, where \(t = 1, \ldots, n_i\), and \(F(x, z) = \Pr(X = x)\), where \(X\) is Poisson distributed with mean \(\exp(z)\). In the standard random-effects model, \(\nu_i\) is assumed to be i.i.d. such that \(\exp(\nu_i)\) is gamma with mean one and variance \(\alpha\), which is estimated from the data. If _normal_ is specified, \(\nu_i\) is assumed to be i.i.d. \(N(0, \sigma^2_{\nu})\).

> Example 1

We have data on the number of ship accidents for five different types of ships (McCullagh and Nelder 1989, 205). We wish to analyze whether the “incident” rate is affected by the period in which the ship was constructed and operated. Our measure of exposure is months of service for the ship, and in this model, we assume that the exponentiated random effects are distributed as gamma with mean one and variance \(\alpha\).
. use https://www.stata-press.com/data/r17/ships
.xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service) irr

Fitting Poisson model:
Iteration 0:  log likelihood = -147.37993
Iteration 1:  log likelihood = -80.372714
Iteration 2:  log likelihood = -80.116093
Iteration 3:  log likelihood = -80.115916
Iteration 4:  log likelihood = -80.115916

Fitting full model:
Iteration 0:  log likelihood = -79.653186
Iteration 1:  log likelihood = -76.990836  (not concave)
Iteration 2:  log likelihood = -74.824942
Iteration 3:  log likelihood = -74.811243
Iteration 4:  log likelihood = -74.811217
Iteration 5:  log likelihood = -74.811217

Random-effects Poisson regression
Number of obs = 34
Group variable: ship
Number of groups = 5
Random effects u_i ~ Gamma

Obs per group:
  min = 6
  avg = 6.8
  max = 7

Wald chi2(4) = 50.90
Log likelihood = -74.811217  Prob > chi2 = 0.0000

|           | IRR  | Std. err. | z     | P>|z|  | [%5% conf. interval] |
|-----------|------|-----------|-------|------|----------------------|
| accident  |      |           |       |      |                      |
| op_75_79  | 1.466305 | .1734005  | 3.24  | 0.001 | 1.162957 1.848777    |
| co_65_69  | 2.032543 | .3040833  | 4.74  | 0.000 | 1.515982 2.72512    |
| co_70_74  | 2.356853 | .3999259  | 5.05  | 0.000 | 1.690033 3.286774   |
| co_75_79  | 1.641913 | .3811398  | 2.14  | 0.033 | 1.04174 2.58786     |
| _cons     | .0013724 | .0002992  | -30.24| 0.000 | .0008952 .002104    |

| ln(service) |      |           |       |      |                      |
|-------------|------|-----------|-------|------|                      |
| lnalpha     | -2.368406 | .8474597 | -4.029397 | -7074155 |
| alpha       | .0936298 | .0793475 | .0177851 | .4929165 |

Note: Estimates are transformed only in the first equation to incidence-rate ratios.
Note: _cons estimates baseline incidence rate (conditional on zero random effects).
LR test of alpha=0: chibar2(01) = 10.61  Prob >= chibar2 = 0.001

The output also includes a likelihood-ratio test of $\alpha = 0$, which compares the panel estimator with the pooled (Poisson) estimator.

We find that the incidence rate for accidents is significantly different for the periods of construction and operation of the ships and that the random-effects model is significantly different from the pooled model.

We may alternatively fit a fixed-effects specification instead of a random-effects specification:
. xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service) irr fe

Iteration 0:  log likelihood = -80.738973
Iteration 1:  log likelihood = -54.857546
Iteration 2:  log likelihood = -54.641897
Iteration 3:  log likelihood = -54.641859
Iteration 4:  log likelihood = -54.641859

Conditional fixed-effects Poisson regression
Group variable: ship

|                        | IRR       | Std. err. | z    | P>|z|   | [95% conf. interval] |
|------------------------|-----------|-----------|------|-------|---------------------|
| accident               |           |           |      |       |                     |
| op_75_79               | 1.468831  | .1737218  | 3.25 | 0.001 | 1.164926            | 1.852019          |
| co_65_69               | 2.008002  | .3004803  | 4.66 | 0.000 | 1.497577            | 2.692398          |
| co_70_74               | 2.26693   | .384865   | 4.82 | 0.000 | 1.625274            | 3.161912          |
| co_75_79               | 1.573695  | .3669393  | 1.94 | 0.052 | .996427            | 2.485397          |
| ln(service)            | 1 (exposure) |         |      |       |                     |                    |

Both of these models fit the same thing but will differ in efficiency, depending on whether the assumptions of the random-effects model are true.
We could have assumed that the random effects followed a normal distribution, $N(0, \sigma^2_\nu)$, instead of a “log-gamma” distribution, and obtained

```
xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service) irr
> normal nolog
```

Random-effects Poisson regression
Group variable: ship
Random effects $u_i \sim$ Gaussian
Integration method: mvaghermite

|              | IRR   | Std. err. | z     | P>|z| | [95% conf. interval] |
|--------------|-------|-----------|-------|------|---------------------|
| accident     |       |           |       |      |                     |
| op_75_79     | 1.466 | .1734     | 3.24  | 0.001| 1.163259 1.849236   |
| co_65_69     | 2.032 | .3040     | 4.74  | 0.000| 1.516025 2.725205   |
| co_70_74     | 2.357 | .3998     | 5.05  | 0.000| 1.690338 3.286717   |
| co_75_79     | 1.647 | .3202     | 2.15  | 0.031| 1.045278 2.594905   |
| _cons        | 0.001 | .0002     | -31.28| 0.000| .0008625 .001982   |
| ln(service)  | 1 (exposure) |       |       |      |                     |
| /lnsig2u     | -2.351| .8586     | -4.034| 0.000| -.6689918          |
| sigma_u      | 0.308 | .1325     | 1.330 | 0.045| .7156988           |

Note: Estimates are transformed only in the first equation to incidence-rate ratios.
Note: _cons estimates baseline incidence rate (conditional on zero random effects).

```
LR test of sigma_u=0: chibar2(01) = 10.67      Prob >= chibar2 = 0.001
```

The output includes the additional panel-level variance component. This is parameterized as the log of the variance $\ln(\sigma^2_\nu)$ (labeled lnsig2u in the output). The standard deviation $\sigma_\nu$ is also included in the output labeled sigma_u.

When $\sigma_\nu$ is zero, the panel-level variance component is unimportant and the panel estimator is no different from the pooled estimator. A likelihood-ratio test of this is included at the bottom of the output. This test formally compares the pooled estimator (poisson) with the panel estimator. Here $\sigma_\nu$ is significantly greater than zero, so a panel estimator is indicated.
Example 2

This time we fit a robust equal-correlation population-averaged model:

```
. xtpoisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service) pa
> vce(robust) eform
```

```
Iteration 1: tolerance = .04083192
Iteration 2: tolerance = .00270188
Iteration 3: tolerance = .00030663
Iteration 4: tolerance = .00003466
Iteration 5: tolerance = 3.891e-06
Iteration 6: tolerance = 4.359e-07
```

```
GEE population-averaged model
Number of obs = 34
Group variable: ship
Number of groups = 5
Family: Poisson
Obs per group:
Link: Log
Correlation: exchangeable
Wald ch2(4) = 252.94
```

```
Scale parameter = 1
Prob > chi2 = 0.0000
(Std. err. adjusted for clustering on ship)
```

```
Robust
| accident  | IRR    | std. err. | z     | P>|z|   | [95% conf. interval] |
|-----------|--------|-----------|-------|-------|---------------------|
| op_75_79  | 1.483299 | .1197901  | 4.88  | 0.000 | 1.266153 1.737685  |
| co_65_69  | 2.038477 | .1809524  | 8.02  | 0.000 | 1.712955 2.425889  |
| co_70_74  | 2.643467 | .4093947  | 6.28  | 0.000 | 1.951407 3.500962  |
| co_75_79  | 1.876656 | .33075    | 3.57  | 0.000 | 1.328511 2.650966  |
| _cons     | .0010255 | .0000721  | -97.90| 0.000 | .0008935 0.001177 |
```

Note: _cons estimates baseline incidence rate (conditional on zero random effects).

We may compare this with a pooled estimator with clustered robust-variance estimates:

```
. poisson accident op_75_79 co_65_69 co_70_74 co_75_79, exp(service)
> vce(cluster ship) irr
```

```
Iteration 0: log pseudolikelihood = -147.37993
Iteration 1: log pseudolikelihood = -80.372714
Iteration 2: log pseudolikelihood = -80.116093
Iteration 3: log pseudolikelihood = -80.115916
Iteration 4: log pseudolikelihood = -80.115916
```

```
Poisson regression
Number of obs = 34
Wald chi2(3) = .
Prob > chi2 = .
Log pseudolikelihood = -80.115916
```

```
(Std. err. adjusted for 5 clusters in ship)
```

```
Robust
| accident  | IRR     | std. err. | z     | P>|z|   | [95% conf. interval] |
|-----------|---------|-----------|-------|-------|---------------------|
| op_75_79  | 1.47324 | .1287036  | 4.44  | 0.000 | 1.2414 1.748377    |
| co_65_69  | 2.125914 | .2850531 | 5.62  | 0.000 | 1.634603 2.764897  |
| co_70_74  | 2.860138 | .6213563 | 4.84  | 0.000 | 1.868384 4.378325  |
| co_75_79  | 2.021926 | .4265285 | 3.34  | 0.000 | 1.337221 3.057227  |
| _cons     | .0009609 | .0000727 | -240.66| 0.000 | .000908 0.0010168 |
```

Note: _cons estimates baseline incidence rate.
Technical note

The random-effects model is calculated using quadrature, which is an approximation whose accuracy depends partially on the number of integration points used. We can use the quadchk command to see if changing the number of integration points affects the results. If the results change, the quadrature approximation is not accurate given the number of integration points. Try increasing the number of integration points using the intpoints() option and run quadchk again. Do not attempt to interpret the results of estimates when the coefficients reported by quadchk differ substantially. See [XT] quadchk for details and [XT] xtprobit for an example.

Because the xtpoisson, re normal likelihood function is calculated by Gauss–Hermite quadrature, on large problems the computations can be slow. Computation time is roughly proportional to the number of points used for the quadrature.

Stored results

xtpoisson, re stores the following in e():

Scalars

\begin{itemize}
  \item e(N) number of observations
  \item e(N_g) number of groups
  \item e(k) number of parameters
  \item e(k_aux) number of auxiliary parameters
  \item e(k_eq) number of equations in e(b)
  \item e(k_eq_model) number of equations in overall model test
  \item e(k_dv) number of dependent variables
  \item e(df_m) model degrees of freedom
  \item e(ll) log likelihood
  \item e(ll_0) log likelihood, constant-only model
  \item e(ll_c) log likelihood, comparison model
  \item e(chi2) \(\chi^2\)
  \item e(chi2_c) \(\chi^2\) for comparison test
  \item e(chi2_c) \(\chi^2\) for comparison test
  \item e(alpha) value of alpha
  \item e(g_min) smallest group size
  \item e(g_avg) average group size
  \item e(g_max) largest group size
  \item e(p) p-value for model test
  \item e(rank) rank of e(V)
  \item e(rank0) rank of e(V) for constant-only model
  \item e(ic) number of iterations
  \item e(rc) return code
  \item e(converged) 1 if converged, 0 otherwise
\end{itemize}

Macros

\begin{itemize}
  \item e(cmd) xtpoisson
  \item e(cmdline) command as typed
  \item e(depvar) name of dependent variable
  \item e(ivar) variable denoting groups
  \item e(model) re
  \item e(wtype) weight type
  \item e(wexp) weight expression
  \item e(title) title in estimation output
  \item e(clustvar) name of cluster variable
  \item e(offset) linear offset variable
  \item e(chi2type) Wald or LR: type of model \(\chi^2\) test
  \item e(chi2_ct) Wald or LR: type of model \(\chi^2\) test corresponding to e(chi2_c)
  \item e(vce) vcetype specified in vce()
In addition to the above, the following is stored in \texttt{r()}: 

Matrices 
\texttt{r(table)} \hspace{1cm} \text{matrix containing the coefficients with their standard errors, test statistics, } p\text{-values, and confidence intervals}

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

\texttt{xtpoisson, re normal} stores the following in \texttt{e()}: 

Scalars 
\texttt{e(N)} \hspace{1cm} \text{number of observations} 
\texttt{e(N_g)} \hspace{1cm} \text{number of groups} 
\texttt{e(k)} \hspace{1cm} \text{number of parameters} 
\texttt{e(k_aux)} \hspace{1cm} \text{number of auxiliary parameters} 
\texttt{e(k_eq)} \hspace{1cm} \text{number of equations in } \texttt{e(b)} 
\texttt{e(k_eq_model)} \hspace{1cm} \text{number of equations in overall model test} 
\texttt{e(k_dv)} \hspace{1cm} \text{number of dependent variables} 
\texttt{e(df_m)} \hspace{1cm} \text{model degrees of freedom} 
\texttt{e(ll)} \hspace{1cm} \text{log likelihood} 
\texttt{e(ll_0)} \hspace{1cm} \text{log likelihood, constant-only model} 
\texttt{e(ll_c)} \hspace{1cm} \text{log likelihood, comparison model} 
\texttt{e(chi2)} \hspace{1cm} \chi^2 
\texttt{e(chi2_c)} \hspace{1cm} \chi^2 \text{ for comparison test} 
\texttt{e(N_clust)} \hspace{1cm} \text{number of clusters} 
\texttt{e(sigma_u)} \hspace{1cm} \text{panel-level standard deviation} 
\texttt{e(n_quad)} \hspace{1cm} \text{number of quadrature points} 
\texttt{e(g_min)} \hspace{1cm} \text{smallest group size} 
\texttt{e(g_avg)} \hspace{1cm} \text{average group size} 
\texttt{e(g_max)} \hspace{1cm} \text{largest group size} 
\texttt{e(p)} \hspace{1cm} \text{ } p\text{-value for model test} 
\texttt{e(rank)} \hspace{1cm} \text{rank of } \texttt{e(V)} 
\texttt{e(rank0)} \hspace{1cm} \text{rank of } \texttt{e(V)} \text{ for constant-only model} 
\texttt{e(ic)} \hspace{1cm} \text{number of iterations} 
\texttt{e(rc)} \hspace{1cm} \text{return code} 
\texttt{e(converged)} \hspace{1cm} 1 \text{ if converged, 0 otherwise} 

Macros 
\texttt{e(cmd)} \hspace{1cm} \text{xtpoisson}
xtpoisson — Fixed-effects, random-effects, and population-averaged Poisson models 15

\[
e(\text{cmdline}) \quad \text{command as typed}
\]
\[
e(\text{depvar}) \quad \text{name of dependent variable}
\]
\[
e(\text{ivar}) \quad \text{variable denoting groups}
\]
\[
e(\text{model}) \quad \text{re}
\]
\[
e(\text{wttype}) \quad \text{weight type}
\]
\[
e(\text{wexp}) \quad \text{weight expression}
\]
\[
e(\text{title}) \quad \text{title in estimation output}
\]
\[
e(\text{clustvar}) \quad \text{name of cluster variable}
\]
\[
e(\text{offset}) \quad \text{linear offset variable}
\]
\[
e(\text{chi2type}) \quad \text{Wald or LR; type of model } \chi^2 \text{ test}
\]
\[
e(\text{chi2_ct}) \quad \text{Wald or LR; type of model } \chi^2 \text{ test corresponding to } e(\text{chi2}\_c)
\]
\[
e(\text{vce}) \quad \text{vcetype specified in } \text{vce}()
\]
\[
e(\text{vcetype}) \quad \text{title used to label Std. err.}
\]
\[
e(\text{intmethod}) \quad \text{integration method}
\]
\[
e(\text{distrib}) \quad \text{Gaussian; the distribution of the random effect}
\]
\[
e(\text{opt}) \quad \text{type of optimization}
\]
\[
e(\text{which}) \quad \text{max or min; whether optimizer is to perform maximization or minimization}
\]
\[
e(\text{ml\_method}) \quad \text{type of ml method}
\]
\[
e(\text{user}) \quad \text{name of likelihood-evaluator program}
\]
\[
e(\text{technique}) \quad \text{maximization technique}
\]
\[
e(\text{properties}) \quad b \ V
\]
\[
e(\text{predict}) \quad \text{program used to implement } \text{predict}
\]
\[
e(\text{marginsdefault}) \quad \text{default } \text{predict() specification for } \text{margins}
\]
\[
e(\text{asbalanced}) \quad \text{factor variables } \text{fvset as asbalanced}
\]
\[
e(\text{asobserved}) \quad \text{factor variables } \text{fvset as asobserved}
\]

Matrices
\[
e(b) \quad \text{coefficient vector}
\]
\[
e(Cns) \quad \text{constraints matrix}
\]
\[
e(\text{ilog}) \quad \text{iteration log}
\]
\[
e(\text{gradient}) \quad \text{gradient vector}
\]
\[
e(V) \quad \text{variance–covariance matrix of the estimators}
\]
\[
e(V\_\text{modelbased}) \quad \text{model-based variance}
\]

Functions
\[
e(\text{sample}) \quad \text{marks estimation sample}
\]

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

Matrices
\[
r(\text{table}) \quad \text{matrix containing the coefficients with their standard errors, test statistics, } p\text{-values,}
\]
\[
\quad \text{and confidence intervals}
\]

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.

\texttt{xtpoisson, fe} stores the following in \( e() \):

Scalars
\[
e(N) \quad \text{number of observations}
\]
\[
e(N\_g) \quad \text{number of groups}
\]
\[
e(k) \quad \text{number of parameters}
\]
\[
e(k\_eq) \quad \text{number of equations in } e(b)
\]
\[
e(k\_eq\_model) \quad \text{number of equations in overall model test}
\]
\[
e(k\_dv) \quad \text{number of dependent variables}
\]
\[
e(df\_m) \quad \text{model degrees of freedom}
\]
\[
e(ll) \quad \text{log likelihood}
\]
\[
e(ll\_c) \quad \text{log likelihood, comparison model}
\]
\[
e(chi2) \quad \chi^2
\]
\[
e(g\_min) \quad \text{smallest group size}
\]
\[
e(g\_avg) \quad \text{average group size}
\]
\[
e(g\_max) \quad \text{largest group size}
\]
\[
e(p) \quad p\text{-value for model test}
\]
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)
xtpoisson
- e(cmdline)
command as typed
- e(depvar)
name of dependent variable
- e(ivar)
variable denoting groups
- e(model)
fe
- e(vtype)
weight type
- e(wtype)
weight expression
- e(title)
title in estimation output
- e(offset)
linear offset variable
- e(chi2type)
Wald; type of model \( \chi^2 \) test
- e(vce)
vcetype specified in vce()
- e(vcetype)
title used to label Std. err.
- e(method)
requested estimation method
- 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(properties)
b V
- e(predict)
program used to implement predict
- 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
- e(gradient)
gradient vector
- e(V)
variance–covariance matrix of the estimators

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.

xtpoisson, pa stores the following in e():

Scalars

- e(N)
number of observations
- e(N_g)
number of groups
- e(df_m)
model degrees of freedom
- e(chi2)
\( \chi^2 \)
- e(p)
\( p \)-value for model test
- e(df_pear)
degrees of freedom for Pearson \( \chi^2 \)
- e(chi2_dev)
\( \chi^2 \) test of deviance
- e(chi2_dis)
\( \chi^2 \) test of deviance dispersion
- e(deviance)
deviance
- e(dispers)
deviance dispersion
- e(phi)
scale parameter
- e(g_min)
smallest group size
- e(g_avg)
average group size
In addition to the above, the following is stored in \( r() \):

Matrices
\[
\begin{align*}
\text{r(table)} & : \text{matrix containing the coefficients with their standard errors, test statistics, } p\text{-values, and confidence intervals}
\end{align*}
\]

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

\texttt{xtpoisson, pa} reports the population-averaged results obtained by using \texttt{xtgee, family(poisson) link(log)} to obtain estimates. See \cite{XT} \texttt{xtgee} for details about the methods and formulas.

\texttt{xtpoisson, fe} with robust standard errors implements the formula presented in \cite{wooldridge99}. The formula is a cluster–robust estimate of the VCE in which the ID variable specifies the clusters.

Although \cite{hausman84} wrote the seminal article on the random-effects and fixed-effects models, \cite{cameron13} provide a good textbook treatment. \cite{allison09, chap. 4} succinctly discusses these models and illustrates the differences between them using Stata.
For a random-effects specification, we know that

\[
Pr(y_{i1}, \ldots, y_{in_i} | \alpha_i, x_{i1}, \ldots, x_{in_i}) = \left( \prod_{t=1}^{n_i} \lambda_{it}^{y_{it}} \right) \exp \left\{ - \exp(\alpha_i) \sum_{t=1}^{n_i} \lambda_{it} \right\} \exp \left( \alpha_i \sum_{t=1}^{n_i} y_{it} \right)
\]

where \( \lambda_{it} = \exp(x_{it} \beta) \). We may rewrite the above as \([\text{defining} \, \epsilon_i = \exp(\alpha_i)]\)

\[
Pr(y_{i1}, \ldots, y_{in_i} | \epsilon_i, x_{i1}, \ldots, x_{in_i}) = \left( \prod_{t=1}^{n_i} \frac{\lambda_{it} \epsilon_i^{y_{it}}}{y_{it}!} \right) \exp \left( \epsilon_i \sum_{t=1}^{n_i} \lambda_{it} \right) \sum_{t=1}^{n_i} y_{it}
\]

We now assume that \( \epsilon_i \) follows a gamma distribution with mean one and variance \( 1/\theta \) so that unconditional on \( \epsilon_i \)

\[
Pr(y_{i1}, \ldots, y_{in_i} | X_i) = \frac{\theta^\theta}{\Gamma(\theta)} \left( \prod_{t=1}^{n_i} \lambda_{it}^{y_{it}} \right) \int_0^\infty \exp \left( -\epsilon_i \sum_{t=1}^{n_i} \lambda_{it} \right) \epsilon_i^{\theta-1} \exp(-\theta \epsilon_i) d\epsilon_i
\]

\[
= \frac{\theta^\theta}{\Gamma(\theta)} \left( \prod_{t=1}^{n_i} \lambda_{it}^{y_{it}} \right) \int_0^\infty \exp \left( -\epsilon_i \left( \theta + \sum_{t=1}^{n_i} \lambda_{it} \right) \right) \epsilon_i^{\theta+\sum_{t=1}^{n_i} y_{it}-1} d\epsilon_i
\]

\[
= \left( \prod_{t=1}^{n_i} \frac{\lambda_{it}^{y_{it}}}{y_{it}!} \right) \frac{\Gamma \left( \theta + \sum_{t=1}^{n_i} y_{it} \right)}{\Gamma(\theta)} \left( \frac{\theta}{\theta + \sum_{t=1}^{n_i} \lambda_{it}} \right)^\theta \left( \frac{1}{\theta + \sum_{t=1}^{n_i} \lambda_{it}} \right) \sum_{t=1}^{n_i} y_{it}
\]

for \( X_i = (x_{i1}, \ldots, x_{in_i}) \).

The log likelihood (assuming gamma heterogeneity) is then derived using

\[
u_i = \frac{\theta}{\theta + \sum_{t=1}^{n_i} \lambda_{it}} \quad \lambda_{it} = \exp(x_{it} \beta)
\]

\[
Pr(Y_{i1} = y_{i1}, \ldots, Y_{in_i} = y_{in_i} | X_i) = \frac{\prod_{t=1}^{n_i} \lambda_{it}^{y_{it}} \Gamma \left( \theta + \sum_{t=1}^{n_i} y_{it} \right)}{\prod_{t=1}^{n_i} y_{it}! \Gamma(\theta) \sum_{t=1}^{n_i} \lambda_{it}} \left( u_i \theta (1 - u_i) \sum_{t=1}^{n_i} y_{it} \right)^\theta
\]
such that the log likelihood may be written as

\[
L = \sum_{i=1}^{n} w_i \left\{ \log \Gamma \left( \theta + \sum_{t=1}^{n_i} y_{it} \right) - \sum_{t=1}^{n_i} \log \Gamma \left( 1 + y_{it} \right) - \log \Gamma(\theta) + \theta \log u_i \\
+ \log(1 - u_i) \sum_{t=1}^{n_i} y_{it} + \sum_{t=1}^{n_i} y_{it} x_{it} \beta - \left( \sum_{t=1}^{n_i} y_{it} \right) \log \left( \sum_{t=1}^{n_i} \lambda_{it} \right) \right\}
\]

where \( w_i \) is the user-specified weight for panel \( i \); if no weights are specified, \( w_i = 1 \).

Alternatively, if we assume a normal distribution, \( N(0, \sigma^2) \), for the random effects \( \nu_i \)

\[
\Pr(y_{i1}, \ldots, y_{in_i}|X_i) = \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \frac{\nu_i^2}{\sigma^2} \right\} \prod_{t=1}^{n_i} F(y_{it}, x_{it} \beta + \nu_i) \, d\nu_i
\]

where

\[
F(y, z) = \exp \left\{ -\exp(z) + yz - \log(y!) \right\}.
\]

The panel-level likelihood \( l_i \) is given by

\[
l_i = \int_{-\infty}^{\infty} \frac{e^{-\nu_i^2/2\sigma^2}}{\sqrt{2\pi\sigma}} \prod_{t=1}^{n_i} F(y_{it}, x_{it} \beta + \nu_i) \, d\nu_i
\]

\[
\equiv \int_{-\infty}^{\infty} g(y_{it}, x_{it}, \nu_i) \, d\nu_i
\]

This integral can be approximated with \( M \)-point Gauss–Hermite quadrature

\[
\int_{-\infty}^{\infty} e^{-x^2} h(x) \, dx \approx \sum_{m=1}^{M} w_m^* h(a_m^*)
\]

This is equivalent to

\[
\int_{-\infty}^{\infty} f(x) \, dx \approx \sum_{m=1}^{M} w_m^* \exp \left\{ (a_m^*)^2 \right\} f(a_m^*)
\]

where the \( w_m^* \) denote the quadrature weights and the \( a_m^* \) denote the quadrature abscessas. The log likelihood, \( L \), is the sum of the logs of the panel-level likelihoods \( l_i \).

The default approximation of the log likelihood is by adaptive Gauss–Hermite quadrature, which approximates the panel-level likelihood with

\[
l_i \approx \sqrt{2\hat{\sigma}_i} \sum_{m=1}^{M} w_m^* \exp \left\{ (a_m^*)^2 \right\} g(y_{it}, x_{it}, \sqrt{2\hat{\sigma}_i} a_m^* + \hat{\mu}_i)
\]
where $\hat{\sigma}_i$ and $\hat{\mu}_i$ are the adaptive parameters for panel $i$. Therefore, with the definition of $g(y_{it}, x_{it}, \nu_i)$, the total log likelihood is approximated by

$$L \approx \sum_{i=1}^{n} w_i \log \left[ \sqrt{2\hat{\sigma}_i} \sum_{m=1}^{M} w_m \exp \left\{ \left( a_m^* \right)^2 \right\} \frac{\exp \left\{ -\frac{\left( \sqrt{2\hat{\sigma}_i} a_m^* + \hat{\mu}_i \right)^2}{2\hat{\sigma}_\nu^2} \right\}}{\sqrt{2\pi\hat{\sigma}_\nu}} \right]$$

$$\prod_{t=1}^{n_i} F(y_{it}, x_{it}\beta + \sqrt{2\hat{\sigma}_i} a_{m_t}^* + \hat{\mu}_i)$$

where $w_i$ is the user-specified weight for panel $i$; if no weights are specified, $w_i = 1$.

The default method of adaptive Gauss–Hermite quadrature is to calculate the posterior mean and variance and use those parameters for $\hat{\mu}_i$ and $\hat{\sigma}_i$ by following the method of Naylor and Smith (1982), further discussed in Skrondal and Rabe-Hesketh (2004). We start with $\hat{\sigma}_{i,0} = 1$ and $\hat{\mu}_{i,0} = 0$, and the posterior means and variances are updated in the $k$th iteration. That is, at the $k$th iteration of the optimization for $l_i$, we use

$$l_{i,k} \approx \sum_{m=1}^{M} \sqrt{2\hat{\sigma}_{i,k-1}} w_m^* \exp \left\{ \left( a_m^* \right)^2 \right\} g(y_{it}, x_{it}, \sqrt{2\hat{\sigma}_{i,k-1}} a_m^* + \hat{\mu}_{i,k-1})$$

Letting

$$\tau_{i,m,k-1} = \sqrt{2\hat{\sigma}_{i,k-1}} a_m^* + \hat{\mu}_{i,k-1}$$

$$\hat{\mu}_{i,k} = \sum_{m=1}^{M} (\tau_{i,m,k-1}) \sqrt{2\hat{\sigma}_{i,k-1}} w_m^* \exp \left\{ \left( a_m^* \right)^2 \right\} g(y_{it}, x_{it}, \tau_{i,m,k-1})$$

and

$$\hat{\sigma}_{i,k} = \sum_{m=1}^{M} (\tau_{i,m,k-1})^2 \sqrt{2\hat{\sigma}_{i,k-1}} w_m^* \exp \left\{ \left( a_m^* \right)^2 \right\} g(y_{it}, x_{it}, \tau_{i,m,k-1})$$

and this is repeated until $\hat{\mu}_{i,k}$ and $\hat{\sigma}_{i,k}$ have converged for this iteration of the maximization algorithm. This adaptation is applied on every iteration until the log-likelihood change from the preceding iteration is less than a relative difference of 1e–6; after this, the quadrature parameters are fixed.

The log likelihood can also be calculated by nonadaptive Gauss–Hermite quadrature, the int-method(ghermite) option, where $\rho = \sigma_\nu^2/(\sigma^2_{\nu} + 1)$:

$$L = \sum_{i=1}^{n} w_i \log \left\{ \Pr(y_{i1}, \ldots, y_{in_i}|x_{i1}, \ldots, x_{in_i}) \right\}$$

$$\approx \sum_{i=1}^{n} w_i \log \left[ \frac{1}{\sqrt{\pi}} \sum_{m=1}^{M} w_m \prod_{t=1}^{n_i} F\left( y_{it}, x_{it}\beta + a_m^* \left( \frac{2\rho}{1 - \rho} \right)^{1/2} \right) \right]$$

Both quadrature formulas require that the integrated function be well approximated by a polynomial of degree equal to the number of quadrature points. The number of periods (panel size) can affect whether

$$\prod_{t=1}^{n_i} F(y_{it}, x_{it}\beta + \nu_i)$$
is well approximated by a polynomial. As panel size and $\rho$ increase, the quadrature approximation can become less accurate. For large $\rho$, the random-effects model can also become unidentified. Adaptive quadrature gives better results for correlated data and large panels than nonadaptive quadrature; however, we recommend that you use the `quadchk` command (see [XT] quadchk) to verify the quadrature approximation used in this command, whichever approximation you choose.

For a fixed-effects specification, we know that

$$
\Pr(Y_{it} = y_{it}|x_{it}) = \exp\{-\exp(\alpha_i + x_{it}\beta)\} \exp(\alpha_i + x_{it}\beta)^{y_{it}}/y_{it}!
$$

$$
= \frac{1}{y_{it}!} \exp\{-\exp(\alpha_i) \exp(x_{it}\beta) + \alpha_i y_{it}\} \exp(x_{it}\beta)^{y_{it}}
$$

$$
\equiv F_{it}
$$

Because we know that the observations are independent, we may write the joint probability for the observations within a panel as

$$
\Pr(Y_{i1} = y_{i1}, \ldots, Y_{in_i} = y_{in_i} | X_i)
$$

$$
= \prod_{t=1}^{n_i} \frac{1}{y_{it}!} \exp\{-\exp(\alpha_i) \exp(x_{it}\beta) + \alpha_i y_{it}\} \exp(x_{it}\beta)^{y_{it}}
$$

$$
= \left(\prod_{t=1}^{n_i} \frac{\exp(x_{it}\beta)^{y_{it}}}{y_{it}!}\right) \exp\{-\exp(\alpha_i) \sum_t \exp(x_{it}\beta) + \alpha_i \sum_t y_{it}\}
$$

and we also know that the sum of $n_i$ Poisson independent random variables, each with parameter $\lambda_{it}$ for $t = 1, \ldots, n_i$, is distributed as Poisson with parameter $\sum_t \lambda_{it}$. Thus

$$
\Pr\left(\sum_t Y_{it} = \sum_t y_{it} | X_i\right) =
$$

$$
\frac{1}{(\sum_t y_{it})!} \exp\left\{-\exp(\alpha_i) \sum_t \exp(x_{it}\beta) + \alpha_i \sum_t y_{it}\right\} \left\{\sum_t \exp(x_{it}\beta)\right\}^{\sum_t y_{it}}
$$

So the conditional likelihood is conditioned on the sum of the outcomes in the set (panel). The appropriate function is given by

$$
\Pr\left(Y_{i1} = y_{i1}, \ldots, Y_{in_i} = y_{in_i}, \sum_t Y_{it} = \sum_t y_{it}\right) =
$$

$$
\left[\left(\prod_{t=1}^{n_i} \frac{\exp(x_{it}\beta)^{y_{it}}}{y_{it}!}\right) \exp\left\{-\exp(\alpha_i) \sum_t \exp(x_{it}\beta) + \alpha_i \sum_t y_{it}\right\}\right] /
$$

$$
\left[\frac{1}{(\sum_t y_{it})!} \exp\left\{-\exp(\alpha_i) \sum_t \exp(x_{it}\beta) + \alpha_i \sum_t y_{it}\right\} \left\{\sum_t \exp(x_{it}\beta)\right\}^{\sum_t y_{it}}\right]
$$

$$
= \left(\sum_t y_{it}\right)! \prod_{t=1}^{n_i} \frac{\exp(x_{it}\beta)^{y_{it}}}{y_{it}! \left\{\sum_k \exp(x_{ik}\beta)\right\}^{y_{it}}}
$$

which is free of $\alpha_i$. 

The conditional log likelihood is given by

\[
L = \log \prod_{i=1}^{n} \left[ \left( \sum_{t=1}^{n_i} y_{it} \right)! \prod_{t=1}^{n_i} \frac{\exp(x_{it}\beta)^{y_{it}}}{\sum_{\ell=1}^{n} \exp(x_{i\ell}\beta)^{y_{i\ell}}^{y_{it}}} \right]^{w_i}
\]

\[
= \log \prod_{i=1}^{n} \left\{ \left( \sum_{t=1}^{n_i} y_{it} \right)! \prod_{t=1}^{n_i} p_{it}^{y_{it}} \right\}^{w_i}
\]

\[
= \sum_{i=1}^{n} w_i \left\{ \log \Gamma \left( \sum_{t=1}^{n_i} y_{it} + 1 \right) - \sum_{t=1}^{n_i} \log \Gamma \left( y_{it} + 1 \right) + \sum_{t=1}^{n_i} y_{it} \log p_{it} \right\}
\]

where

\[
p_{it} = \frac{e^{x_{it}\beta}}{\sum_{\ell} e^{x_{i\ell}\beta}}
\]

**xtpoisson, re and the robust VCE estimator**

Specifying vce(robust) or vce(cluster clustvar) causes the Huber/White/sandwich VCE estimator to be calculated for the coefficients estimated in this regression. See [P] _robust, particularly Introduction and Methods and formulas. Wooldridge (2020) and Arellano (2003) discuss this application of the Huber/White/sandwich VCE estimator. As discussed by Wooldridge (2020), Stock and Watson (2008), and Arellano (2003), specifying vce(robust) is equivalent to specifying vce(cluster panelvar), where panelvar is the variable that identifies the panels.

Clustering on the panel variable produces a consistent VCE estimator when the disturbances are not identically distributed over the panels or there is serial correlation in $\epsilon_{it}$.

The cluster–robust VCE estimator requires that there are many clusters and the disturbances are uncorrelated across the clusters. The panel variable must be nested within the cluster variable because of the within-panel correlation that is generally induced by the random-effects transform when there is heteroskedasticity or within-panel serial correlation in the idiosyncratic errors.

**References**


Also see

[XT] `xtpoisson postestimation` — Postestimation tools for xtpoisson

[XT] `quadchk` — Check sensitivity of quadrature approximation

[XT] `xtgee` — Fit population-averaged panel-data models by using GEE

[XT] `xtnbreg` — Fixed-effects, random-effects, & population-averaged negative binomial models

[XT] `xtset` — Declare data to be panel data

[BAYES] `bayes: xtpoisson` — Bayesian random-effects Poisson model

[ME] `mepoisson` — Multilevel mixed-effects Poisson regression

[M] `Estimation` — Estimation commands for use with mi estimate

[R] `poisson` — Poisson regression

[U] 20 Estimation and postestimation commands