Package 'chantrics'

October 12, 2022

Type Package

Title Loglikelihood Adjustments for Econometric Models

Version 1.0.0

Date 2021-09-28

Description Adjusts the loglikelihood of common econometric models for
clustered data based on the estimation process suggested in
Chandler and Bate (2007) <doi:10.1093/biomet/asm015>, using the 'chandwich'
package <https://cran.r-project.org/package=chandwich>, and provides
convenience functions for inference on the adjusted models.

Imports AER, chandwich, graphics, sandwich, stats, utils, methods,
lmtest, rlang, progress, purrr

Depends R (>= 4.0.0)

URL <https://chantrics.theobruckbauer.eu>,

   <https://github.com/tbruckbauer/chantrics>

BugReports <https://github.com/tbruckbauer/chantrics/issues>

License EUPL (>= 1.2)

Encoding UTF-8

RoxygenNote 7.1.2

Language en-GB

Suggests knitr, rmarkdown, testthat, lax, covr, MASS, pscl

VignetteBuilder knitr

NeedsCompilation no

Author Theo Bruckbauer [aut, cre] (<https://orcid.org/0000-0002-4764-9953>)

Maintainer Theo Bruckbauer <chantrics@theobruckbauer.eu>

Repository CRAN

Date/Publication 2021-09-29 08:20:11 UTC

R topics documented:

---

adj_loglik   *Loglikelihood adjustments for fitted models*

Description

This function adjusts the loglikelihood of fitted model objects based on Chandler and Bate (2007). It is a generic function for different types of models, which are listed in **Supported models**. This section also contains links to function-specific help pages.

Usage

```
adj_loglik(x, cluster = NULL, use_vcov = TRUE, use_mle = TRUE, ...)
```

Arguments

x   A supported fitted model object, see **Supported models**

cluster   A vector or factor indicating the cluster the corresponding loglikelihood contri-
                  bution belongs to. It is required to have the same length as the vector returned
                  by [logLik_vec()](). If cluster is not supplied or NULL, then it is assumed that
                  each observation forms its own cluster.

use_vcov   A logical scalar. By default, the [vcov()]() method for x is used to estimate
                  the Hessian of the independence loglikelihood, if the function exists. Other-
                  wise, or if use_vcov = FALSE, H is estimated using [stats::optimHess()]() inside
                  [chandwich::adjust_loglik()]().

use_mle   A logical scalar. By default, the MLE from x is taken as given, and is not
                  reestimated. By setting use_mle to FALSE, the parameters are reestimated in the
                  function [chandwich::adjust_loglik()]() using [stats::optim()]().This feature
                  is currently for development purposes only, may return misleading/false results
                  and may be removed without notice.

...   Further arguments to be passed to [sandwich::meatCL()]() if cluster is defined,
                  if cluster = NULL, they are passed into [sandwich::meat()]().

Details

If use_vcov = TRUE, the current default, the function will test whether a vcov S3 method exists for x, and will take the variance-covariance matrix from there. Otherwise, or if use_vcov = FALSE the variance-covariance matrix of the MLE is estimated inside chandwich::adjust_loglik() using stats::optimHess().

Value

An object of class "chantrics" inheriting from class "chandwich". See the documentation provided with chandwich::adjust_loglik().

Supported models

- glm
- glm.nb

Available methods

"chantrics" objects have the following methods available to them:

- AIC
- anova
- alrtest - Adjusted Likelihood ratio tests
- coef
- lmtest::coeftest - (z) tests for all coefficients
- confint and plot.confint - confidence intervals for all coefficients, and diagnostics plots for confint().
- conf_intervals - enhanced confidence interval reports
- conf_region - two-dimensional confidence regions
- df.residual
- fitted
- logLik
- logLik_vec
- nobs
- plot
- predict
- print
- residuals
- summary
- terms
- vcov

Examples

See the model-specific pages in the *supported models* section.

References

R. E. Chandler and S. Bate, Inference for clustered data using the independence loglikelihood, Biometrika, 94 (2007), pp. 167–183. doi: [10.1093/biomet/asm015](#).

See Also

[lax::alogLik()](#) supports adjustment for user-supplied objects.

---

alrtest	*Adjusted Likelihood Ratio Test of Nested Models*

Description

alrtest is a helper function to simulate the functions [lmtest::waldtest()](#) and [lmtest::lrtest()](#) for adjusted chantrics objects. The method can be employed to compare nested models (see details).

Usage

```
alrtest(object, ...)
```

Arguments

object   a chantrics object as returned from [adj_loglik()](#).

...   further object specifications (see details), as well as named parameters that will be passed to [chandwich::compare_models()](#). The type of adjustment, out of "vertical", "cholesky", "spectral", "none", as specified in the parameter type, can also be specified here.

Details

This function is a helper function that creates an interface to [anova.chantrics()](#) that is similar to [lmtest::waldtest()](#) and [lmtest::lrtest()](#).

The standard method is to compare the fitted model object object with the models in .... Instead of passing the fitted models into ..., other specifications are possible. Note that the types of specifications cannot be mixed, except between numerics/characters. The type of the second object supplied determines the algorithm used.

- "chantrics" **objects**: When supplying two or more "chantrics" objects, they will be sorted as in [anova.chantrics()](#). Then, the ALRTS will be computed consecutively between the two neighbouring models. Note that all models must be nested. For details refer to [anova.chantrics()](#).

- "numeric": If the second object is "numeric" or "character", then "numeric" objects corresponding element in attr(terms(object1), "term.labels") will be turned into their corresponding "character" element and will be handled as in "character" below.

- "character": If the second object is "numeric" or "character", then the "character" objects are consecutively included in an update formula like update(object1, . ~ . - object2)

- "formula": If the second object is a "formula", then the second model will be computed as update(object1, object2).

Then, the adjusted likelihood ratio test statistic (ALRTS), as described in Section 3.5 of Chandler and Bate (2007), is computed by [anova.chantrics()](#).

If a single unnamed object is passed in ..., sequential ANOVA is performed on object.

Value

An object of class "anova" inheriting from class "data.frame". The columns are as follows:

Resid.df	The residual number of degrees of freedom in the model.
df	The increase in residual degrees of freedom with respect to the model in the row above.
ALRTS	The adjusted likelihood ratio statistic.
Pr(>ALRTS)	The p-value of the test that the model above is a "significantly better" model as the one in the current row.

References

R. E. Chandler and S. Bate, Inference for clustered data using the independence loglikelihood, Biometrika, 94 (2007), pp. 167–183. doi: [10.1093/biomet/asm015](#).

See Also

[anova.chantrics()](#) for the implementation of the computations of the test statistics.

[lmtest::waldtest()](#) and [lmtest::lrtest()](#) for syntax.

---

anova.chantrics   *ANOVA tables: compare nested models*

Description

anova method for chantrics objects

Usage

```
## S3 method for class 'chantrics'
anova(object, ...)
```

Arguments

object	Object of class `chantrics`, as returned by `adj_loglik()`.
...	Further objects of class `chantrics`, as returned by `adj_loglik()`, and named parameters that should be passed to `chandwich::compare_models()`. The type of adjustment, out of `"vertical"`, `"cholesky"`, `"spectral"`, `"none"`, as specified in the parameter `type`, can also be specified here.

Details

Create an analysis of adjusted deviance table for one object (sequential), or two or more nested models that have been adjusted using the `adj_loglik()` method. It uses the adjusted likelihood ratio test statistic (ALRTS), as described in Section 3.5 of Chandler and Bate (2007).

Each line represents the model as given above the table, with each line (except for the first line) showing the residual degrees of freedom of that model, the change in degrees of freedom, the ALRTS and the associated p-value in comparison to the model in the line above.

When a single model is specified, the function returns a sequential analysis of deviance table, where, iteratively, one term is being removed from the right of the full formula. This process is continued until the "intercept only" model is left. The row names are the names of the dropped term in comparison to the model in the line above.

If more than one model is specified, the function sorts the models by their number of variables as returned by `adj_loglik()` in `attr(x, "p_current")`.

Details of the ALRT can be found in `chandwich::compare_models()` and in Chandler and Bate (2007).

Value

An object of class `"anova"` inheriting from class `"data.frame"`. The columns are as follows:

Resid.df	The residual number of degrees of freedom in the model.
df	The increase in residual degrees of freedom with respect to the model in the row above.
ALRTS	The adjusted likelihood ratio statistic.
Pr(>ALRTS)	The p-value of the test that the model above is a "significantly better" model as the one in the current row.

References

R. E. Chandler and S. Bate, Inference for clustered data using the independence loglikelihood, Biometrika, 94 (2007), pp. 167–183. doi: 10.1093/biomet/asm015.

See Also

chandwich::compare_models: implementation of the comparison mechanism

Examples

```
# from Introducing Chandwich.
set.seed(123)
x <- rnorm(250)
y <- rnbinom(250, mu = exp(1 + x), size = 1)
fm_pois <- glm(y ~ x + I(x^2), family = poisson)
fm_pois_adj <- adj_loglik(fm_pois)
fm_pois_small_adj <- update(fm_pois_adj, formula = . ~ . - I(x^2))
fm_pois_smallest_adj <- update(fm_pois_adj, formula = . ~ 1)

anova(fm_pois_adj, fm_pois_small_adj, fm_pois_smallest_adj)
# use different types of adjustment with type, default is "vertical"
anova(fm_pois_adj, fm_pois_small_adj, fm_pois_smallest_adj, type = "cholesky")

# sequential anova
anova(fm_pois_adj)
```

---

chantrics   *chantrics: Loglikelihood Adjustments for Econometric Models*

Description

`chantrics` adjusts the loglikelihood of common econometric models for clustered data based on the estimation process suggested in Chandler and Bate (2007), using the chandwich package, and provides convenience functions for inference on the adjusted models. `adj_loglik()` adjusts the model's parameter covariance matrix to incorporate clustered data, and can mitigate model misspecification by wrapping `chandwich::adjust_loglik` for the supported models.

Details

The returned model of class `chantrics` can be plugged into standard model evaluation and model comparison methods, for example, `summary()`, `confint()` and `anova()`, and a hypothesis test framework provided by `alrtest()`.

See `vignette("chantrics-vignette", package = "chantrics")` for an overview of the package.

References

R. E. Chandler and S. Bate, Inference for clustered data using the independence loglikelihood, Biometrika, 94 (2007), pp. 167–183. doi: 10.1093/biomet/asm015.

---

glm   *Loglikelihood adjustments for glm fits*

Description

In a generalised linear model (glm), the user can choose between a range of distributions of a response y, and can allow for non-linear relations between the mean outcome for a particular combination of covariates x, $E(y_i \mid x_i) = \mu_i$, and the linear predictor, $\eta_i = x_i^T \beta$, which is the link function $g(\mu_i) = \eta_i$. it is required to be monotonic. (For a quick introduction, see Kleiber and Zeileis (2008, Ch. 5.1), for more complete coverage of the topic, see, for example, Davison (2003, Ch. 10.3))

Details

For more usage examples and more information on `glm` models, see the *Introducing* `chantrics` vignette by running `vignette("chantrics-vignette", package = "chantrics")`

Supported families (within each family, any link function should work)

- `gaussian`
- `poisson`
- `binomial`
- `MASS::negative.binomial`

Also works for `MASS::glm.nb()`, note that the standard errors of the theta are not adjusted.

References

Davison, A. C. 2003. Statistical Models. Cambridge Series on Statistical and Probabilistic Mathematics 11. Cambridge University Press, Cambridge.

Kleiber, Christian, and Achim Zeileis. 2008. Applied Econometrics with R. Edited by Robert Gentleman, Kurt Hornik, and Giovanni Parmigiani. Use r! New York: Springer-Verlag.

Examples

```
# binomial example from Applied Econometrics in R, Kleiber/Zeileis (2008)
# ==  probit  ==
data("SwissLabor", package = "AER")
swiss_probit <- glm(participation ~ . + I(age^2),
  data = SwissLabor,
  family = binomial(link = "probit")
)
summary(swiss_probit)
swiss_probit_adj <- adj_loglik(swiss_probit)
summary(swiss_probit_adj)

# == logit ==
swiss_logit <- glm(participation ~ . + I(age^2),
```

```
  data = SwissLabor,
  family = binomial(link = "logit")
)
summary(swiss_logit)
swiss_logit_adj <- adj_loglik(swiss_logit)
summary(swiss_logit_adj)
```

---

logLik_vec   *Evaluate loglikelihood contributions from specific observations*

Description

Generic function for calculating the loglikelihood contributions from individual observations for a fitted model.

Usage

```
logLik_vec(object, ...)
```

Arguments

object	A fitted model object.
...	Further arguments.

Value

An object of class "logLik_vec", which is a numeric vector of length nobs(object) (i.e. the number of observations in object) of the loglikelihood of each observation. Additionally, it contains the attributes df (model degrees of freedom) and nobs (number of observations).

The methods [stats::logLik()](), and [stats::nobs()]() are available.

See Also

[stats::logLik()]()

---

predict.chantrics   *Predict Method for chantrics fits*

Description

Obtains predictions from chantrics objects. The function can currently only supply predictions of the link and the response values of the data used for the fit.

Usage

```
## S3 method for class 'chantrics'
predict(object, newdata = NULL, type = c("response", "link"), ...)
```

Arguments

object	Object of class chantrics, as returned by adj_loglik()
newdata	optionally, a data frame in which to look for variables with which to predict. If omitted, the fitted linear predictors are used. Supplying new data is currently not supported.
type	the type of prediction required. The default "response" is on the scale of the response variables. The alternative "link" is on the scale of the linear predictors, if applicable. Otherwise, an error is returned.
...	unused.

Details

If newdata is omitted, the predictions are based on the data used for the fit. Any instances of NA will return NA.

Value

A vector of predictions.

---

residuals.chantrics   *Residuals of chantrics model fits*

Description

residuals() returns the residuals specified in type from a "chantrics" object.

Usage

```
## S3 method for class 'chantrics'
residuals(object, type = c("response", "working", "pearson"), ...)
```

Arguments

object	an object of class "chantrics", returned by adj_loglik().
type	the type of residuals which should be returned. The alternatives are: "response" (default), "working", and "pearson" (for glm fits).
...	further arguments passed to or from other methods

Details

The different types of residuals are as in stats::residuals.glm().

Value

A vector of residuals.

References

A. C. Davison and E. J. Snell, Residuals and diagnostics. In: Statistical Theory and Modelling. In Honour of Sir David Cox, FRS, 1991. Eds. Hinkley, D. V., Reid, N. and Snell, E. J., Chapman & Hall.

M. Döring, Interpreting Generalised Linear Models. In: Data Science Blog, 2018. `https://www.datascienceblog.net/post/machine-learning/interpreting_generalized_linear_models/`

See Also

`adj_loglik()` for model fitting, `stats::residuals.glm()`, and `stats::residuals()`.

---

update.chantrics   *Update, re-fit and re-adjust a Model Call*

Description

update.chantrics() will update a model that has been adjusted by `adj_loglik()`. It passes all arguments to the standard `stats::update()` function.

Usage

```
## S3 method for class 'chantrics'
update(object, ...)
```

Arguments

object	A "chantrics" returned by `adj_loglik()`.
...	Additional arguments to the call, passed to `stats::update()` to update the original model specification.

Details

The function cannot change any arguments passed to the adj_loglik() function. To change any of these arguments, re-run adj_loglik().

Passing evaluate = FALSE is not supported, if this is required, run `stats::update()` on the unadjusted object.

Value

The fitted, adjusted "chantrics" object.

See Also

`stats::update()`

`stats::update.formula()`

Examples

```
# from Introducing Chandwich.
set.seed(123)
x <- rnorm(250)
y <- rnbinom(250, mu = exp(1 + x), size = 1)
fm_pois <- glm(y ~ x + I(x^2), family = poisson)
fm_pois_adj <- adj_loglik(fm_pois)
fm_pois_small_adj <- update(fm_pois_adj, formula = . ~ . - I(x^2))
summary(fm_pois_small_adj)
fm_pois_smallest_adj <- update(fm_pois_adj, formula = . ~ 1)
summary(fm_pois_smallest_adj)
```

Index