

Package ‘quantification’

November 12, 2016

Type Package

Title Quantification of Qualitative Survey Data

Version 0.2.0

Date 2016-10-30

Author Joachim Zuckarelli <joachim@zuckarelli.de>

Maintainer Joachim Zuckarelli <joachim@zuckarelli.de>

Description Provides different functions for quantifying qualitative survey data. It supports the Carlson-Parkin method, the regression approach, the balance approach and the conditional expectations method.

License GPL-3

Repository CRAN

Depends car, stats

LazyLoad yes

NeedsCompilation no

Date/Publication 2016-11-12 00:13:24

R topics documented:

quantification-package	2
bal	3
ce	6
cp	9
ra	13

Index	18
--------------	-----------

quantification-package

Quantification of qualitative survey data

Description

This package provides different functions for quantifying qualitative survey data. It supports the Carlson-Parkin method (developed by Carlson/Parkin (1985)), the regression approach (developed by Pesaran (1984)), the balance approach (see Batchelor (1984)) and the conditional expectations method.

These methods are usually applied in economics to quantify qualitative inflation expectations collected through surveys (e.g. consumer surveys). However, the same approaches can be used to quantify *any* qualitative expectation about the change of a variable that is quantitative in nature.

The four functions of the package – cp for the Carlson-Parkin method, ra for the regression approach, bal for the balance approach and ce for the conditional expectations method – allow the user to customize a wide range of parameters and make use of certain extensions of the original methods. Apart from that all the functions deliver per default two versions of quantified expectations: one under the assumption that the survey respondents form expectations over the *absolute* change of the variable in question and one under the assumption of expectations over the *relative* change. The functions also provide the user with standard measures for the forecast quality of the quantified expectations enabling the user to quickly assess the effect of a change in the quantification method used.

Comments and suggestions

Your comments and suggestions are highly appreciated.

Author(s)

Joachim Zuckarelli

Maintainer: Joachim Zuckarelli <joachim@zuckarelli.de>

References

- Batchelor, R.A. (1984), Quantitative vs. qualitative measures of inflation expectations, *Oxford Bulletin of Economics and Statistics* **48** (2), 99–120.
- Carlson, J. A./Parkin, M. (1975), Inflation expectations, *Economica* **42**, 123–138.
- Henzel, S./Wollmershaeuser, T. (2005), Quantifying inflation expectations with the Carlson-Parkin method: A survey-based determination of the just noticeable difference, *Journal of Business Cycle Measurement and Analysis* **2**, 321–352.
- Nardo, M. (2003), The quantification of qualitative survey data: a critical assessment, *Journal of Economic Surveys* **17** (5), 645–668.
- Pesaran, M. (1984), Expectations formation and macroeconomic modelling, in: Malgrange, M. (1984), *Contemporary macroeconomic modelling*, 27–55.

Zuckarelli, J. (2015): A new method for quantification of qualitative expectations, *Economics and Business Letters* 3(5), Special Issue *Energy demand forecasting*, 123-128.

bal *The balance approach*

Description

bal implements the balance approach for the quantification of qualitative survey data. A description of the method can be found in Batchelor (1986).

Usage

```
bal(y.series, survey.up, survey.same, survey.down, forecast.horizon,
    first.period = 1, last.period = (length(survey.up) - forecast.horizon),
    growth.limit = NA, suppress.warnings = FALSE)
```

Arguments

y.series	a numerical vector containing the variable whose change is the subject of the qualitative survey question. If, for example the survey asks participants to assess whether inflation will increase, decrease or stay the same, y.series would be the series of inflation data.
survey.up	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to increase. This vector needs to be of the same length as y.series.
survey.same	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to stay the same. This vector needs to be of the same length as y.series.
survey.down	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to decrease. This vector needs to be of the same length as y.series.
forecast.horizon	a numeric value defining the number of periods the survey question looks in to the future. If the data in y.series is monthly data and the survey question asks respondents to assess the development of the variable over the next six months then forecast.horizon=6.
first.period	an optional numeric value indexing the first period for which survey data in survey.up, survey.same and survey.down shall be used for quantification; default value is 1.
last.period	an optional numeric value indexing the last period for which survey data in survey.up, survey.same and survey.down shall be used for quantification; default value is length(survey.up) - forecast.horizon.

- `growth.limit` serves to limit the effect of outliers when expectations are quantified under the assumption that survey respondents form expectations on the percentage change of y . `growth.limit` defines a limit for percentage change of y . When this limit is exceeded the growth rate is set automatically to the median growth of y over the period covered by the expectations. Default value is NA.
- `suppress.warnings` a logical value indicating if runtime warnings shall be displayed (FALSE) or not (TRUE). Default value is FALSE.

Details

`bal` provides two alternative versions of quantified expectations, depending on the assumed expectation formation process of survey respondents. The basic common assumption of the balance method is that survey participants are asked to assess whether variable y will go up or down or stay the same. Survey respondents can now form expectations on either the absolute or the relative change of y which differ because the scaling factors (thetas) used to scale the difference between the shares of 'up' and the 'down' respondents are calculated differently in each case. The `bal` function calculates both versions.

The survey result vectors `survey.up`, `survey.down` and `survey.same` as well as the variable `y.series` must be of the same length and must cover the forecasted horizon (i.e. `last.period + forecast.horizon ≤ length(survey.up)`).

Data in `survey.up`, `survey.down` and `survey.same` outside the survey period interval [`first.period`, `last.period`] are ignored. Similarly, `y.series` data with a period index greater than `last.period` is ignored.

`survey.up`, `survey.down` and `survey.same` need not sum up to 100% or 1 (which may happen, for example, if the survey has a 'Don't know' answer option).

Value

- `y.e.mean.abs`: a numeric vector containing the quantified mean expectations of the variable y , assuming that survey respondents form expectations on the *absolute* change in y . For all periods which are not in scope of the survey the value is NA.
- `y.e.mean.perc`: a numeric vector containing the quantified mean expectations of the variable y , assuming that survey respondents form expectations on the *relative* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.abs`: a numeric vector containing the quantified mean absolute change of the variable y , assuming that survey respondents form expectations on the *absolute* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.perc`: a numeric vector containing the quantified mean percentage change of the variable y , assuming that survey respondents form expectations on the *relative* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.sd.abs`: a numeric vector containing the standard deviation of the absolute change expectation for variable y in the population. Assumes that survey respondents form expectations on the *absolute* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.sd.perc`: a numeric vector containing the standard deviation of the absolute change expectation for variable y in the population. Assumes that survey respondents form

expectations on the *relative* change in y . For all periods which are not in scope of the survey the value is NA.

- `theta.abs`: a numeric vector containing the estimated factor which scales the difference between the shares of 'up' respondents and 'down' respondents assuming that survey respondents form expectations on the *absolute* change in variable y . For all periods which are not in scope of the survey the value is NA.
- `theta.perc`: a numeric vector containing the estimated factor which scales the difference between the shares of 'up' respondents and 'down' respondents assuming that survey respondents form expectations on the *relative* change in variable y . For all periods which are not in scope of the survey the value is NA.
- `nob`: a numeric value showing the number of periods for which expectations have been quantified.
- `mae.abs`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *absolute* change in variable y .
- `rmse.abs`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *absolute* change in variable y .
- `mae.perc`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *relative* change in variable y .
- `rmse.perc`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *relative* change in variable y .

Please cite as:

Zuckarelli, Joachim (2014). *Quantification of qualitative survey data in R*.
R package version 1.0.0. <http://CRAN.R-project.org/package=quantification>

Author(s)

Joachim Zuckarelli, <joachim@zuckarelli.de>

References

Batchelor, R.A. (1984), Quantitative vs. qualitative measures of inflation expectations, *Oxford Bulletin of Economics and Statistics* **48** (2), 99–120.

See Also

[quantification-package](#), [cp](#), [ra](#), [ce](#)

Examples

```
## Data preparation: generate a sample dataset with inflation and survey data
inflation<-c(1.7, 1.9, 2, 1.9, 2, 2.1, 2.1, 2.1, 2.4, 2.3, 2.4)
answer.up<-c(67, 75.1, 76.4, 72.4, 69.7, 49.7, 45.2, 31.6, 14.9, 19.3, 19.2)
answer.same<-c(30.1, 19.6, 19.5, 21.3, 20.1, 33.1, 34.4, 33.5, 44.6, 38.1, 35.3)
answer.down<-c(2.9, 5.3, 4.1, 6.3, 10.2, 17.2, 20.4, 34.9, 40.5, 42.6, 45.5)

## Call bal for quantification
```

```
quant.bal<-bal(inflation, answer.up, answer.same, answer.down, first.period=5,
last.period=7, forecast.horizon=4)
```

ce

*Conditional Expectations method***Description**

ce implements the Conditional Expectations approach for the quantification of qualitative survey data. The method calculates expectations on a distribution of past realizations of the variable of interest (variable y), conditional on the expectation of either an increase or a decrease in y . These conditional expectations are then weighted with the share of survey respondents expecting variable y to rise or fall, respectively. For details see

Usage

```
ce(y.series, survey.up, survey.same, survey.down, forecast.horizon,
first.period = 11, last.period = (length(survey.up) - forecast.horizon),
exp.horizon.type = "moving", mov.horizon.length = 10,
fix.horizon.start = 1, fix.horizon.end = 10,
distrib.param = "mean", suppress.warnings = FALSE)
```

Arguments

- | | |
|-------------------------------|--|
| <code>y.series</code> | a numerical vector containing the variable whose change is the subject of the qualitative survey question. If, for example the survey asks participants to assess whether inflation will increase, decrease or stay the same, <code>y.series</code> would be the series of inflation data. |
| <code>survey.up</code> | a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to increase. This vector needs to be of the same length as <code>y.series</code> . |
| <code>survey.same</code> | a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to stay the same. This vector needs to be of the same length as <code>y.series</code> . |
| <code>survey.down</code> | a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to decrease. This vector needs to be of the same length as <code>y.series</code> . |
| <code>forecast.horizon</code> | a numeric value defining the number of periods the survey question looks in to the future. If the data in <code>y.series</code> is monthly data and the survey question asks respondents to assess the development of the variable over the next six months then <code>forecast.horizon=6</code> . |
| <code>first.period</code> | an optional numeric value indexing the first period for which survey data in <code>survey.up</code> , <code>survey.same</code> and <code>survey.down</code> shall be used for quantification; default value is 11. |

<code>last.period</code>	an optional numeric value indexing the last period for which survey data in <code>survey.up</code> , <code>survey.same</code> and <code>survey.down</code> shall be used for quantification; default value is <code>length(survey.up) - forecast.horizon</code> .
<code>exp.horizon.type</code>	<p>an optional character vector indicating the type of experience horizon to be used. The experience horizon is the time period over which the distribution of variable <code>y</code> serves as basis for quantification. Possible values are:</p> <ul style="list-style-type: none"> • "fix": the distribution of variable <code>y</code> over a fixed period of time is used for calculating the conditional expectations • "moving": the time period over which the distribution of variable <code>y</code> is used for quantification moves with the time the respective survey results were generated, i.e. the experience horizon reaches back a certain number of periods from the time of the survey. <p>Default value is "moving".</p>
<code>mov.horizon.length</code>	an optional numeric value indicating the length of the (moving) forecast horizon. Is only considered when <code>exp.horizon.type="moving"</code> . Default value is 10.
<code>fix.horizon.start</code>	an optional numeric value indicating the first period of the (fixed) forecast horizon. Is only considered when <code>exp.horizon.type="fix"</code> . Default value is 1.
<code>fix.horizon.end</code>	an optional numeric value indicating the last period of the (fixed) forecast horizon. Is only considered when <code>exp.horizon.type="fix"</code> . Default value is 10.
<code>distrib.param</code>	an optional character vector indicating the distribution parameter that shall be used for calculating conditional expectations based on the distribution of variable <code>y</code> over the expectation horizon. Can be either "mean" or "median". Default value is "mean".
<code>suppress.warnings</code>	a logical value indicating if runtime warnings shall be displayed (FALSE) or not (TRUE). Default value is FALSE.

Details

The survey result vectors `survey.up`, `survey.down` and `survey.same` as well as the variable `y.series` must be of the same length and must cover the forecasted horizon (i.e. `last.period + forecast.horizon ≤ length(survey.up)`).

Data in `survey.up`, `survey.down` and `survey.same` outside the survey period interval [`first.period`, `last.period`] are ignored. Similarly, `y.series` data with a period index greater than `last.period` is ignored.

`survey.up`, `survey.down` and `survey.same` need not sum up to 100% or 1 (which may happen, for example, if the survey has a 'Don't know' answer option).

Value

`ce` returns a list containing the quantified survey data and some meta information. The list has the following elements:

- `y.e`: a numeric vector containing the quantified expectations of the variable `y`.

- nob: a numeric value showing the number of periods for which expectations have been quantified.
- mae: a numeric value showing the mean absolute error (MAE) of expectations.
- rmse: a numeric value showing the root mean squared error (RMSE) of expectations.

Please cite as:

Zuckarelli, Joachim (2014). *Quantification of qualitative survey data in R*.
R package version 1.0.0. <http://CRAN.R-project.org/package=quantification>

Author(s)

Joachim Zuckarelli, <joachim@zuckarelli.de>

References

Zuckarelli, J. (2015): A new method for quantification of qualitative expectations, *Economics and Business Letters* 3(5), Special Issue *Energy demand forecasting*, 123-128.

See Also

[quantification-package](#), [cp](#), [bal](#), [ra](#)

Examples

```
## Data preparation: generate a sample dataset with inflation and survey data
inflation<-c(1.5, 1.5, 1.5, 1.1, 0.9, 1.3, 1.3, 1.2, 1.7, 1.7, 1.5, 2, 1.4, 1.9, 1.9, 2.3, 2.8,
2.5, 2.1, 2.1, 1.9, 1.9, 1.5, 1.6, 2.1, 1.8, 2.1, 1.5, 1.3, 1.1, 1.1, 1.3, 1.3, 1.3, 1.1,
1.1, 1, 1.2, 1.1, 0.9)
answer.up<-c(72.7, 69.7, 60.9, 53.7, 54.9, 54.8, 56.1, 51.7, 62.2, 54.2, 39.8, 18.6, 5.4, 8.2,
8.6, 8.5, 16, 18.9, 7.7, 6.5, 6.4, 7, 7.4, 6.8, 9.5, 17.1, 13.1, 21.5, 22.7, 26.9, 32.4,
20.2, 20.4, 15.8, 11.4, 7.9, 11.3, 10, 11.3, 9.7)
answer.same<-c(24.1, 22.8, 24.3, 26.2, 31.1, 35.4, 33, 35.5, 27.4, 24.8, 32.1, 44.8, 41.8,
37.9, 33.2, 30.9, 29.9, 22.1, 17.2, 15.5, 21.8, 25.2, 23.2, 24.2, 32.9, 31.2, 42.2, 50.5,
52.5, 56.3, 53.8, 62.8, 65.6, 63, 60.3, 61.1, 57.8, 63, 61.4, 61.9)
answer.down<-c(3.2, 7.5, 14.8, 20.1, 14, 9.8, 10.9, 12.8, 10.4, 21, 28.1, 36.6, 52.8, 53.9,
58.2, 60.6, 54.1, 59, 75.1, 78, 71.8, 67.8, 69.4, 69, 57.6, 51.7, 44.7, 28, 24.8, 16.8,
13.8, 17, 14, 21.2, 28.3, 31, 30.9, 27, 27.3, 28.4)

## Call ce for quantification
quant.ce<-ce(inflation, answer.up, answer.same, answer.down, first.period=30, last.period=36,
forecast.horizon=4, exp.horizon.type = "fix", fix.horizon.start = 1, fix.horizon.end = 29)
```

cp

*Carlson-Parkin method***Description**

cp implements the method for the quantification of qualitative survey data proposed by Carlson/Parkin (1985). Additionally, it provides certain extensions of the Carlson-Parkin approach (e.g. other distributions than the normal distribution, indifference limens depending on the level of the forecasted variable).

Usage

```
cp(y.series, survey.up, survey.same, survey.down, forecast.horizon,
  first.period = 1, last.period = (length(survey.up) - forecast.horizon),
  limen.type = "carlson.parkin", const.limen = 0, user.symm.limen = 0,
  user.upper.limen = 0, user.lower.limen = 0, correct.zero = TRUE,
  correct.by = 0.01, growth.limit = NA, distrib.type = "normal",
  distrib.mean = 0, distrib.sd = 1, distrib.log.location = 0,
  distrib.log.scale = 1, distrib.t.df = (last.period - first.period),
  suppress.warnings = FALSE)
```

Arguments

y.series	a numerical vector containing the variable whose change is the subject of the qualitative survey question. If, for example the survey asks participants to assess whether inflation will increase, decrease or stay the same, y.series would be the series of inflation data.
survey.up	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to increase. This vector needs to be of the same length as y.series.
survey.same	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to stay the same. This vector needs to be of the same length as y.series.
survey.down	a numerical vector containing the number or the share of survey respondents expecting the variable contained in y.series to decrease. This vector needs to be of the same length as y.series.
forecast.horizon	a numeric value defining the number of periods the survey question looks in to the future. If the data in y.series is monthly data and the survey question asks respondents to assess the development of the variable over the next six months then forecast.horizon=6.
first.period	an optional numeric value indexing the first period for which survey data in survey.up, survey.same and survey.down shall be used for quantification; default value is 1.

<code>last.period</code>	an optional numeric value indexing the last period for which survey data in <code>survey.up</code> , <code>survey.same</code> and <code>survey.down</code> shall be used for quantification; default value is <code>length(survey.up) - forecast.horizon</code> .
<code>limen.type</code>	an optional character vector describing the type of indifference limen that shall be used for quantification. Possible values are: <ul style="list-style-type: none"> • <code>"carlson.parkin"</code>: standard symmetric, time-invariant Carlson-Parkin indifference limens; default value. • <code>"weber.fechner"</code>: symmetric, time-varying indifference limens following the Weber-Fechner law, i.e. proportional to the variable in question (i.e. <code>y.series</code>) at the time of the survey. For details see Henzel/Wollmershaeuser (2005). • <code>"constant"</code>: symmetric, time-invariant indifference limen defined by the user through the <code>const.limen</code> argument • <code>"symm.series"</code>: fully user-defined, symmetric, time-varying indifference limen given by the argument <code>user.symm.limen</code>. • <code>"asymm.series"</code>: fully user-defined, asymmetric, time-varying indifference limens given by the arguments <code>user.upper.limen</code> and <code>user.lower.limen</code>.
<code>const.limen</code>	an optional numeric value containing the symmetric, time-invariant user-defined indifference limen. Must be provided when <code>"constant"</code> is chosen as <code>limen.type</code> . Is ignored otherwise. Default value is <code>0</code> .
<code>user.symm.limen</code>	an optional numeric vector containing the symmetric, time-varying indifference limen. Must be provided when <code>"symm.series"</code> is chosen as <code>limen.type</code> . Is ignored otherwise. Default value is a <code>0</code> .
<code>user.upper.limen</code>	an optional numeric vector containing the upper, time-varying indifference limen. Must be provided when <code>"asymm.series"</code> is chosen as <code>limen.type</code> . Is ignored otherwise. Default value is <code>0</code> .
<code>user.lower.limen</code>	an optional numeric vector containing the lower, time-varying indifference limen. Must be provided when <code>"asymm.series"</code> is chosen as <code>limen.type</code> . Is ignored otherwise. Default value is <code>0</code> .
<code>correct.zero</code>	an optional logical value steering the automatic correction of zero-values in <code>survey.up</code> and <code>survey.down</code> . If <code>survey.up</code> or <code>survey.down</code> are zero in any period then the Carlson-Parkin method delivers counter-intuitive results (see Nardo (2003) for details). Automatic correction changes the zero value to an amount specified by the <code>correct.by</code> argument. <code>correct.zero=TRUE</code> switches automatic correction on which is the default.
<code>correct.by</code>	an optional parameter indicating the amount by which <code>survey.up</code> and <code>survey.down</code> are corrected in case they have zero value and <code>correct.zero=TRUE</code> . Default value is <code>0.01</code> .
<code>growth.limit</code>	serves to limit the effect of outliers when expectations are quantified under the assumption that survey respondents form expectations on the percentage change of <code>y</code> . <code>growth.limit</code> defines a limit for percentage change of <code>y</code> . When this limit is exceeded the growth rate is set automatically to the median growth of <code>y</code> over the period covered by the expectations. Default value is <code>NA</code> .

<code>distrib.type</code>	an optional character vector describing the type of distribution used for quantification. Possible values are: <ul style="list-style-type: none"> • "normal": the normal distribution is used. Default value for <code>distrib.type</code>. Parameters <code>distrib.mean</code> and <code>distrib.sd</code> can be used to specify the distribution. • "logistic": the logistic distribution is used. Parameters <code>distrib.log.location</code> and <code>distrib.log.scale</code> can be used to specify the distribution. • "t": the t distribution is used. Parameter <code>distrib.t.df</code> can be used to specify the distribution.
<code>distrib.mean</code>	an optional numerical value defining the mean of the normal distribution (used in case <code>distrib.type="normal"</code>). Default value is 0.
<code>distrib.sd</code>	an optional numerical value defining the standard deviation of the normal distribution (used in case <code>distrib.type="normal"</code>). Default value is 1.
<code>distrib.log.location</code>	an optional numerical value defining the location of the logistic distribution (used in case <code>distrib.type="logistic"</code>). Default value is 0.
<code>distrib.log.scale</code>	an optional numerical value defining the scale of the logistic distribution (used in case <code>distrib.type="logistic"</code>). Default value is 1.
<code>distrib.t.df</code>	an optional numerical value defining the degrees of freedom (df) of the t distribution (used in case <code>distrib.type="t"</code>). Default value is <code>last.period - first.period</code> .
<code>suppress.warnings</code>	a logical value indicating if runtime warnings shall be displayed (FALSE) or not (TRUE). Default value is FALSE.

Details

cp provides two alternative versions of quantified expectations, depending on the assumed expectation formation process of survey respondents. The basic common assumption of the Carlson-Parkin method is that survey participants are asked to assess whether variable `y` will go up or down or stay the same. Survey respondents can now form expectations on either the absolute or the relative change of `y` which differ because the indifferent limens used for quantification are calculated differently in each case. The `cp` function calculates both versions.

The survey result vectors `survey.up`, `survey.down` and `survey.same` as well as the variable `y.series` must be of the same length and must cover the forecasted horizon (i.e. `last.period + forecast.horizon ≤ length(survey.up)`).

Data in `survey.up`, `survey.down` and `survey.same` outside the survey period interval [`first.period`, `last.period`] are ignored. Similarly, `y.series` data with a period index greater than `last.period` is ignored.

`first.period` must be greater than `forecast.horizon`, because indifference limens use the current change of variable `y` for calibration. In order to calculate the change in `y` for the survey observation with index `first.period` the observation of `y` with index `first.period - forecast.horizon` is required.

`survey.up`, `survey.down` and `survey.same` need not sum up to 100% or 1 (which may happen, for example, if the survey has a 'Don't know' answer option).

The Weber-Fechner option (see Henzel/Wollmershaeuser (2005) for details) weighs the 'expectations' term in the traditional Carlson-Parkin limen calculation with the current inflation rate. The

resulting time-invariant value is applied as a proportionality factor to current inflation leading to a time-varying, inflation-proportional indifference limen.

When the indifference limen is user-defined (i.e. `limen.type="symm.series"` or `limen.type="asymm.series"`) then each limen value needs to be placed in such a way in the limen vector(s) (i.e. `user.limen` or `user.upper.limen / user.lower.limen`) that it has the same index as the survey observation (in `survey.up`, `survey.down` and `survey.same`) to which it belongs.

Value

`cp` returns a list containing the quantified survey data and some meta information. The list has the following elements:

- `y.e.mean.abs`: a numeric vector containing the quantified mean expectations of the variable `y`, assuming that survey respondents form expectations on the *absolute* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `y.e.mean.perc`: a numeric vector containing the quantified mean expectations of the variable `y`, assuming that survey respondents form expectations on the *relative* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.abs`: a numeric vector containing the quantified mean absolute change of the variable `y`, assuming that survey respondents form expectations on the *absolute* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.perc`: a numeric vector containing the quantified mean percentage change of the variable `y`, assuming that survey respondents form expectations on the *relative* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.sd.abs`: a numeric vector containing the standard deviation of the absolute change expectation for variable `y` in the population. Assumes that survey respondents form expectations on the *absolute* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.sd.perc`: a numeric vector containing the standard deviation of the absolute change expectation for variable `y` in the population. Assumes that survey respondents form expectations on the *relative* change in `y`. For all periods which are not in scope of the survey the value is NA.
- `limen.abs`: a numeric vector containing the estimated (or user-defined) indifference limens assuming that survey respondents form expectations on the *absolute* change in variable `y`. For all periods which are not in scope of the survey the value is NA.
- `limen.perc`: a numeric vector containing the estimated (or user-defined) indifference limens assuming that survey respondents form expectations on the *relative* change in variable `y`. For all periods which are not in scope of the survey the value is NA.
- `nob`: a numeric value showing the number of periods for which expectations have been quantified.
- `mae.abs`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *absolute* change in variable `y`.
- `rmse.abs`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *absolute* change in variable `y`.
- `mae.perc`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *relative* change in variable `y`.

- `rmse.perc`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *relative* change in variable `y`.

Please cite as:

Zuckarelli, Joachim (2014). *Quantification of qualitative survey data in R*.
R package version 0.1.0. <http://CRAN.R-project.org/package=quantification>

Author(s)

Joachim Zuckarelli, <joachim@zuckarelli.de>

References

- Carlson, J. A./Parkin, M. (1975), Inflation expectations, *Economica* **42**, 123–138.
- Henzel, S./Wollmershaeuser, T. (2005), Quantifying inflation expectations with the Carlson-Parkin method: A survey-based determination of the just noticeable difference, *Journal of Business Cycle Measurement and Analysis* **2**, 321–352.
- Nardo, M. (2003), The quantification of qualitative survey data: a critical assessment, *Journal of Economic Surveys* **17** (5), 645–668.

See Also

[quantification-package](#), [bal](#), [ra](#), [ce](#)

Examples

```
## Data preparation: generate a sample dataset with inflation and survey data
inflation<-c(1.7, 1.9, 2, 1.9, 2, 2.1, 2.1, 2.1, 2.4, 2.3, 2.4)
answer.up<-c(67, 75.1, 76.4, 72.4, 69.7, 49.7, 45.2, 31.6, 14.9, 19.3, 19.2)
answer.same<-c(30.1, 19.6, 19.5, 21.3, 20.1, 33.1, 34.4, 33.5, 44.6, 38.1, 35.3)
answer.down<-c(2.9, 5.3, 4.1, 6.3, 10.2, 17.2, 20.4, 34.9, 40.5, 42.6, 45.5)

## Call cp for quantification
quant.cp.limens<-cp(inflation, answer.up, answer.same, answer.down, first.period=5,
last.period=7, forecast.horizon=4)

## With Weber-Fechner limens instead of Carson-Parkin limens
quant.wf.limens<-cp(inflation, answer.up, answer.same, answer.same, first.period=5,
last.period=7, forecast.horizon=4, limen.type="weber.fechner")
```

ra

Regression approach

Description

`ra` implements the regression approach developed by Pesaran (1984).

Usage

```
ra(y.series, survey.up, survey.same, survey.down, forecast.horizon,
  first.period = 1, last.period = (length(survey.up) - forecast.horizon),
  distrib.type = "normal", distrib.mean = 0, distrib.sd = 1,
  distrib.log.location = 0, distrib.log.scale = 1,
  distrib.t.df = (first.period - last.period), growth.limit = NA,
  symmetry.error = "white", suppress.warnings = FALSE)
```

Arguments

<code>y.series</code>	a numerical vector containing the variable whose change is the subject of the qualitative survey question. If, for example the survey asks participants to assess whether inflation will increase, decrease or stay the same, <code>y.series</code> would be the series of inflation data.
<code>survey.up</code>	a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to increase. This vector needs to be of the same length as <code>y.series</code> .
<code>survey.same</code>	a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to stay the same. This vector needs to be of the same length as <code>y.series</code> .
<code>survey.down</code>	a numerical vector containing the number or the share of survey respondents expecting the variable contained in <code>y.series</code> to decrease. This vector needs to be of the same length as <code>y.series</code> .
<code>forecast.horizon</code>	a numeric value defining the number of periods the survey question looks in to the future. If the data in <code>y.series</code> is monthly data and the survey question asks respondents to assess the development of the variable over the next six months then <code>forecast.horizon=6</code> .
<code>first.period</code>	an optional numeric value indexing the first period for which survey data in <code>survey.up</code> , <code>survey.same</code> and <code>survey.down</code> shall be used for quantification; default value is 1.
<code>last.period</code>	an optional numeric value indexing the last period for which survey data in <code>survey.up</code> , <code>survey.same</code> and <code>survey.down</code> shall be used for quantification; default value is <code>length(survey.up) - forecast.horizon</code> .
<code>distrib.type</code>	an optional character vector describing the type of distribution used for quantification. Possible values are: <ul style="list-style-type: none"> • "normal": the normal distribution is used. Default value for <code>distrib.type</code>. Parameters <code>distrib.mean</code> and <code>distrib.sd</code> can be used to specify the distribution. • "logistic": the logistic distribution is used. Parameters <code>distrib.log.location</code> and <code>distrib.log.scale</code> can be used to specify the distribution. • "t": the t distribution is used. Parameter <code>distrib.t.df</code> can be used to specify the distribution.
<code>distrib.mean</code>	an optional numerical value defining the mean of the normal distribution (used in case <code>distrib.type="normal"</code>). Default value is 0.

<code>distrib.sd</code>	an optional numerical value defining the standard deviation of the normal distribution (used in case <code>distrib.type="normal"</code>). Default value is 1.
<code>distrib.log.location</code>	an optional numerical value defining the location of the logistic distribution (used in case <code>distrib.type="logistic"</code>). Default value is 0.
<code>distrib.log.scale</code>	an optional numerical value defining the scale of the logistic distribution (used in case <code>distrib.type="logistic"</code>). Default value is 1.
<code>distrib.t.df</code>	an optional numerical value defining the degrees of freedom (df) of the t distribution (used in case <code>distrib.type="t"</code>). Default value is <code>last.period - first.period</code> .
<code>growth.limit</code>	serves to limit the effect of outliers when expectations are quantified under the assumption that survey respondents form expectations on the percentage change of <i>y</i> . <code>growth.limit</code> defines a limit for percentage change of <i>y</i> . When this limit is exceeded the growth rate is set automatically to the median growth of <i>y</i> over the period covered by the expectations. Default value is NA.
<code>symmetry.error</code>	an optional character vector indicating the type of standard error used for testing the symmetry of the estimated upper and lower indifference limens. Can be either "white" (for White standard error) or "small.sample" (for small sample standard error, HC3), see MacKinnon/White (1985) for details. Default value is "white".
<code>suppress.warnings</code>	a logical value indicating if runtime warnings shall be displayed (FALSE) or not (TRUE). Default value is FALSE.

Details

`ra` estimates the time-invariant, asymmetric indifference limens using OLS regression with non-robust standard errors.

The function `ra` provides two alternative versions of quantified expectations, depending on the assumed expectation formation process of survey respondents. The basic common assumption of the regression approach is that survey participants are asked to assess whether variable *y* will go up or down or stay the same. Survey respondents can now form expectations on either the absolute or the relative change of *y*. The `reg` function calculates both versions.

The survey result vectors `survey.up`, `survey.down` and `survey.same` as well as the variable `y.series` must be of the same length and must cover the forecasted horizon (i.e. `last.period + forecast.horizon ≤ length(survey.up)`).

Data in `survey.up`, `survey.down` and `survey.same` outside the survey period interval [`first.period`, `last.period`] are ignored. Similarly, `y.series` data with a period index greater than `last.period` is ignored.

`survey.up`, `survey.down` and `survey.same` need not sum up to 100% or 1 (which may happen, for example, if the survey has a 'Don't know' answer option).

Value

- `y.e.mean.abs`: a numeric vector containing the quantified mean expectations of the variable *y*, assuming that survey respondents form expectations on the *absolute* change in *y*. For all periods which are not in scope of the survey the value is NA.

- `y.e.mean.perc`: a numeric vector containing the quantified mean expectations of the variable y , assuming that survey respondents form expectations on the *relative* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.abs`: a numeric vector containing the quantified mean absolute change of the variable y , assuming that survey respondents form expectations on the *absolute* change in y . For all periods which are not in scope of the survey the value is NA.
- `delta.y.e.mean.perc`: a numeric vector containing the quantified mean percentage change of the variable y , assuming that survey respondents form expectations on the *relative* change in y . For all periods which are not in scope of the survey the value is NA.
- `upper.limit.abs`: a numeric value containing the estimated upper indifference limen when survey respondents form expectations on the *absolute* change in variable y .
- `lower.limit.abs`: a numeric value containing the estimated upper indifference limen when survey respondents form expectations on the *absolute* change in variable y .
- `upper.limit.perc`: a numeric value containing the estimated upper indifference limen when survey respondents form expectations on the *relative* change in variable y .
- `lower.limit.perc`: a numeric value containing the estimated upper indifference limen when survey respondents form expectations on the *relative* change in variable y .
- `nob`: a numeric value showing the number of periods for which expectations have been quantified.
- `mae.abs`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *absolute* change in variable y .
- `rmse.abs`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *absolute* change in variable y .
- `mae.perc`: a numeric value showing the mean absolute error (MAE) of expectations when survey respondents form expectations on the *relative* change in variable y .
- `rmse.perc`: a numeric value showing the root mean squared error (RMSE) of expectations when survey respondents form expectations on the *relative* change in variable y .
- `symmetry.abs`: a numeric value containing the p-value of the test for symmetry of the estimated indifference limens when survey respondents form expectations on the *absolute* change in variable y . The standard error used for testing depends on the argument `symmetry.error`.
- `symmetry.perc`: a numeric value containing the p-value of the test for symmetry of the estimated indifference limens when survey respondents form expectations on the *relative* change in variable y . The standard error used for testing depends on the argument `symmetry.error`.

Please cite as:

Zuckarelli, Joachim (2014). *Quantification of qualitative survey data in R*.
R package version 1.0.0. <http://CRAN.R-project.org/package=quantification>

Author(s)

Joachim Zuckarelli, <joachim@zuckarelli.de>

References

- MacKinnon, J.G./White, H. (1985), Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties, *Journal of Econometrics* **29**, 305–325.
- Pesaran, M. (1984), Expectations formation and macroeconomic modelling, in: Malgrange, M. (1984), *Contemporary macroeconomic modelling*, 27–55.

See Also

[quantification-package](#), [cp](#), [bal](#), [ce](#)

Examples

```
## Data preparation: generate a sample dataset with inflation and survey data
inflation<-c(1.7, 1.9, 2, 1.9, 2, 2.1, 2.1, 2.1, 2.4, 2.3, 2.4)
answer.up<-c(67, 75.1, 76.4, 72.4, 69.7, 49.7, 45.2, 31.6, 14.9, 19.3, 19.2)
answer.same<-c(30.1, 19.6, 19.5, 21.3, 20.1, 33.1, 34.4, 33.5, 44.6, 38.1, 35.3)
answer.down<-c(2.9, 5.3, 4.1, 6.3, 10.2, 17.2, 20.4, 34.9, 40.5, 42.6, 45.5)

## Call ra for quantification
quant.ra<-ra(inflation, answer.up, answer.same, answer.down, first.period=5,
last.period=7, forecast.horizon=4, symmetry.error="small.sample")
```

Index

[bal](#), [3](#), [8](#), [13](#), [17](#)

[ce](#), [5](#), [6](#), [13](#), [17](#)

[cp](#), [5](#), [8](#), [9](#), [17](#)

quantification

[\(quantification-package\)](#), [2](#)

[quantification-package](#), [2](#)

[ra](#), [5](#), [8](#), [13](#), [13](#)