Package: mmbcv (via r-universe)

May 31, 2026

Title Multistate Model Bias-Corrected Robust Variance

Version 0.3.0

Description Computes robust and bias-corrected sandwich variance estimators for multi-state Cox models with clustered time-to-event data. The methodology extends the marginal Cox model bias-correction framework of Wang et al. (2023) <doi:10.1002/bimj.202200113> to the multi-state setting.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Suggests knitr, rmarkdown, survival, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Can Meng [aut, cre], Denise Esserman [aut], Fan Li [aut], Erich Greene [aut]

Maintainer Can Meng <`can.meng@yale.edu`>

Depends R (>= 3.5.0)

Repository https://mengcan47.r-universe.dev

Date/Publication 2026-03-31 17:18:22 UTC

RemoteUrl https://github.com/cran/mmbcv

RemoteRef HEAD

RemoteSha 7ba5e94a594996f57f6e6962f37db506b20836b7

Contents

MMBCV *MMBCV: Bias-corrected cluster-robust sandwich variance for multistate Cox models*

Description

Computes several sandwich variance estimators (robust, MR, MD, FG, KC, MBN and combinations) for multistate Cox models using cluster-level score aggregation, with Breslow or Efron ties.

Usage

```
MMBCV(
  fit,
  data,
  StartTime,
  StopTime,
  ClusterID,
  SubjectID,
  Event,
  tie = c("breslow", "efron"),
  details = FALSE,
  detail_names = NULL,
  kc_tol = 1e-10
)
```

Arguments

`fit`	A fitted multistate Cox model object with components `cmap`, `rmap`, `states`, and `coefficients`. Typically produced by workflows in the **survival** package.
`data`	A data.frame in counting-process form that matches the fit.
`StartTime`, `StopTime`	Unquoted column names for start/stop times.
`ClusterID`	Unquoted column name for the cluster identifier.
`SubjectID`	Unquoted column name for the subject identifier.
`Event`	Unquoted column name for the event/state indicator used to define transition-specific events.
`tie`	Tie-handling method: `"breslow"` or `"efron"`.
`details`	Logical; if TRUE, returns additional intermediate quantities.
`detail_names`	Optional character vector specifying which extra components of the full sandwich object to append when `details = TRUE`.
`kc_tol`	Numeric tolerance used in the KC positive-semidefinite adjustment.

Value

A list containing variance-covariance matrices:

robust Standard cluster-robust sandwich variance estimator.

varMR Martingale-residual (MR) bias-corrected variance estimator.

varMD Mancl-DeRouen (MD) bias-corrected variance estimator.

varMDMR Hybrid Mancl-DeRouen plus martingale-residual (MDMR) bias-corrected variance estimator.

varFG Fay-Graubard (FG) bias-corrected variance estimator.

varFGMR Hybrid Fay-Graubard plus martingale-residual (FGMR) bias-corrected variance estimator.

varKC Kauermann-Carroll (KC) bias-corrected variance estimator.

varKCMR Hybrid Kauermann-Carroll plus martingale-residual (KCMR) bias-corrected variance estimator.

varMBN Morel-Bokossa-Neerchal (MBN) bias-corrected variance estimator.

varMBNMR Hybrid Morel-Bokossa-Neerchal plus martingale-residual (MBNMR) bias-corrected variance estimator.

If `details = TRUE`, additional intermediate matrices may also be returned.

References

Wang, X., Turner, E. L., and Li, F. (2023). Improving sandwich variance estimation for marginal Cox analysis of cluster randomized trials. *Biometrical Journal*, 65, 2200113.

Schaubel, D. E. (2005). Adjusting analytic inference for small samples in recurrent event regression. *Statistics in Medicine*, 24(19), 3037–3051.

Kauermann, G., and Carroll, R. J. (2001). A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*, 96(456), 1387–1396.

Fay, M. P., and Graubard, B. I. (2001). Small-sample adjustments for Wald-type tests using sandwich estimators. *Biometrics*, 57(4), 1198–1206.

Mancl, L. A., and DeRouen, T. A. (2001). A covariance estimator for GEE with improved small-sample properties. *Biometrics*, 57(1), 126–134.

Morel, J. G., Bokossa, M., and Neerchal, N. K. (2003). Small sample correction for the variance of GEE estimators. *Biometrical Journal*, 45(4), 395–409.

Therneau, T., Crowson, C., and Atkinson, E. (2026). *Multi-state models and competing risks*. Vignette for the **survival** package. See `vignette("compete", package = "survival")` or `https://CRAN.R-project.org/package=survival`.

See Also

`coxph`, `Surv`

Examples

```
data("msdat3")

# Example requires the suggested package 'survival'
if (requireNamespace("survival", quietly = TRUE)) {
  fit <- survival::coxph(
    list(
      survival::Surv(Tstart, Tstop, event) ~ 1,
      state("(s0)"):state("S1") + state("S1"):state("S2") + state("S2"):state("S3") ~
        Z + X,
      state("(s0)"):state("D") + state("S1"):state("D") + state("S2"):state("D") +
        state("S3"):state("D") ~ (Z + X)/common
    ),
    data = msdat3, id = id, ties = "breslow",
    timefix = FALSE
  )

  MMBCV(
       fit       = fit,
       data      = msdat3,
       StartTime = Tstart,
       StopTime  = Tstop,
       ClusterID = clus_id,
       SubjectID = id,
       Event     = event,
       tie       = "breslow",
       details   = FALSE)
}
```

`msdat3` *Clustered multistate simulated dataset (counting-process format)*

Description

A simulated clustered, subject-level multistate dataset in long (counting-process) format. Treatment is assigned at the cluster level and baseline covariates are available at both the cluster and individual levels. Each subject may contribute multiple rows, corresponding to transitions between states over time.

Usage

```
data(msdat3)
```

Format

A data frame with 3,004 rows and 9 variables:

id Subject identifier. There are 1,500 unique subjects.

Z Cluster-level treatment assignment indicator. `Z = 1` indicates the cluster is assigned to treatment; `Z = 0` indicates control.

X Individual-level baseline covariate generated from a standard normal distribution.

clus_id Cluster identifier. There are 30 clusters.

from Starting state for the current observation interval (e.g., `(s0)`).

to Ending state for the current observation interval (e.g., `S1`, `S2`, `S3`, `D`, or `censor`).

Tstart Start time of the interval.

Tstop Stop time of the interval.

event End state at `Tstop`. Values include `censor`, `S1`, `S2`, `S3`, and `D`, where `D` is the absorbing state.

Details

The multistate process uses the following state labels:

- `(s0)`: initial (starting) state.
- `S1`: first intermediate transition state.
- `S2`: second intermediate transition state.
- `S3`: third intermediate transition state.
- `D`: absorbing (terminal) state.

Each row represents a time interval `[Tstart, Tstop]` during which the subject is at risk for the transition indicated by `from` and `to`. The variable `event` records the state at which the interval ends. `event = "censor"` indicates the observation is right-censored at `Tstop` (i.e., no transition is observed at the end of the interval).

The distribution of `event` in this dataset is:

- `censor`: 1050
- `S1`: 837
- `S2`: 450
- `S3`: 217
- `D`: 450

Source

Simulated data bundled with the package for examples and testing.

Examples

```
data("msdat3")
nrow(msdat3)
length(unique(msdat3$id))
length(unique(msdat3$clus_id))
table(msdat3$event)
```

`subset_by_transition` *Subset a dataset by transition using a row map*

Description

Subset a dataset by transition using a row map

Usage

```
subset_by_transition(dataA, dataB, transition)
```

Arguments

`dataA`	Full dataset to subset.
`dataB`	Two-column object: column 1 = row indices in `dataA`, column 2 = transition id.
`transition`	Character vector of transition names used as output list names.

Value

A named list of data frames, one per transition.

Index