Response variables

```
 -27  -13   28   56   84  168  259  331  427  504
 561  334  157  374  191  465  125  212  232  177
 672  771  834  945 1008 1092 1289 1306 1351
  98  207  127  202  143  174  216  245  237

 -14   -6   58  253  358  508  574  672  855  924
 429  587  446  269  131   50  145  634  273  144

 -14   -7    0   56   84  168  336  420  504  672
 231  312  123  127  297  337  225  312  178  111
 756  840  924 1000 1135 1260 1280 1337
  97  133  239  151  115  297  141  113

----------

 5 13  0
12  4  2  0
23  0
 3  3  2  4 14  4  0
 3 13  7  1
 3  6 12  2  0
```

Time-varying covariates

0	28	58	85	113	159	203	333	375	585
1.0	1.167	1.333	1.5	1.667	1.833	2.0	1.833	2.0	0.0
591	1306								
2.0	0.0								

0	218	312	352	403	406	973
1.0	1.2	1.4	1.6	1.8	2.0	0.0

0	29	57	87	119	164	203	241	287	818	835	1280
1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	0.0	2.4	0.0

Time-constant covariates

Each has one value per subject.

Characteristics of a response variable

- Times (if longitudinal).
- Censor indicators (if durations).
- Binomial denominators.
- Unit of measurement (if continuous).
- Has response been transformed? (Jacobian)
- Nesting or clustering indicators.
- Weights (??).

Questions

- How to read in such data?

- How to store them?

- What to do with missing values?

Recording and reading

- Record each type of variable in a separate file. (maximum 3 files)

- If not in balanced rectangular form, read as a list with one element per individual.

Storage

- For modelling efficiency:
  - Lists cannot easily be transferred to C or Fortran.
  - Data for all individuals stored together as vectors, or matrices when several covariates of the same kind.
  - Minimize copying by using R scoping rules.
  - Factor variables?? (bit-encoding?)

- Three classes, corresponding to types of variables.

  - response,

  - tvcov,

  - tccov.

- ‘Response’ class contains all supplied information as separate vectors. (maximum 6)

- ‘Response’ class has supplementary vector indicating number of observations per individual. (7th)

- Most recent time-varying covariate value can be brought forward to response time. (Are ties effective immediately?)

- How to handle missing values?
  - Missing value process virtually never independent of the process of interest.
  - Number of recorded observations can change with the covariates present in the model.
  - Individuals can disappear when model changed.
- Supplementary class, ‘repeated’, for a given combination of variables in a model, with NAs removed.

Methods

- Transform matrices and lists to create objects.

- Print summary information for each type of object, not the whole data array.

- Plot longitudinal responses and time-varying covariates.
  - Choice of subsets.
  - Individual points or profiles.
  - If nesting, times may start at zero in each cluster (e.g. cross-over design).

- Find all information about each individual for a given model:
  - Carry forward time-varying covariates to response times.
  - Create interactions among time-varying covariates or with time-constant covariates.
  - If necessary, for constructing the likelihood, match time-constant covariates to individual responses.
- NA removal:
  - Only possible after all information has been joined for a given model.
  - When applicable, must be coordinated for each given combination of variables.