'rmap' Simulated Data (v.01)

Gail Gong and David Johnston

May 5, 2011

The functions `df_randomSample` and `df_randomSample_r1_r2` generate risk data according to simple random sampling; `df_twoStage` and `df_twoStage_r1_r2` supply data for two-stage sampling. In `df_randomSample` and `df_twoStage`, the column `r` contains the true probability of disease. In `df_randomSample_r1_r2` and `df_twoStage_r1_r2` the column `r` is replaced with `r1` and a second risk `r2` is given; `r2` is a noisy version of `r1`.

This document contains details for the way `df_randomSample` generates data.

Individuals in the data set are subject to times (t0) to censoring, times (t1) to disease, and time to death (t2). We observe t = min(t0, t1, t2) and e = which of t0, t1, t2 is equal to t.

The generating functions generate t0, t1, and t2 independently according to exponential distributions with rates = eta0, eta1, and eta2 respectively, where eta0 and eta2 are inputs and eta1 is generated from either a lognormal or beta distribution. If lognormal, then log(t1) is normal with mean param1 and sd param2. If beta, then param1 and param2 are the shape parameters of beta.

Given these models for generating data, the probability of developing disease given a person's eta1 can be calculated using equation 18 of in "rmap-formulas-v01.pdf" from the rmap website. This probability is calculated and recorded as r in the output variables.

```
library(rmap)
options(width = 100)
options(digits = 4)
set.seed(1)
df1 = df_randomSample()
head(df1)
```

```
     e      t      w      r c k
 446 0 6.0090 0.1287 0.5055 A 2
 975 1 1.9604 0.1779 0.6004 A 3
 656 0 1.6690 0.1183 0.4809 A 2
 232 1 3.5440 0.3129 0.7456 A 5
 799 0 0.5533 0.1886 0.6170 A 4
 361 1 2.9491 0.1191 0.4827 A 2
```

```
ddd = df_twoStage()
df2 = ddd$d
N = ddd$N
N
n = ddd$n
head(df2)
```

```
   A   B
 439 561
     e     t       w      r c k
 292 2 3.410 0.09281 0.4114 B 1
 474 1 1.325 0.24635 0.6890 A 5
 821 1 2.675 0.21790 0.6569 A 4
 947 0 4.258 0.11646 0.4763 B 1
 462 1 1.162 0.12879 0.5058 A 2
 590 0 1.061 0.13756 0.5252 B 2
```