Package ‘KMsurv’

February 19, 2015

Version 0.1-5
Date 2012/12/03
Title Data sets from Klein and Moeschberger (1997), Survival Analysis
Author Original by Klein and Moeschberger, modifications by Jun Yan
    <jun.yan@uconn.edu>
Maintainer Jun Yan <jun.yan@uconn.edu>
Description Data sets and functions for Klein and Moeschberger (1997),
    "Survival Analysis, Techniques for Censored and Truncated
    Data", Springer.
License GPL (>= 3)
Repository CRAN
Date/Publication 2012-12-03 15:51:22
NeedsCompilation no

R topics documented:

aids ................................................................. 2
alloauto ............................................................ 3
allograft ............................................................ 3
azt ................................................................. 4
baboon ............................................................. 4
bcdeter ........................................................... 5
bfeed ............................................................... 5
bmt ............................................................... 6
bnct ............................................................. 7
btrial ............................................................ 7
burn .............................................................. 8
channing ......................................................... 9
drug6mp ........................................................ 9
drughiv .......................................................... 10
hodg ............................................................. 10
kidney ........................................................... 11
The `aids` data frame has 295 rows and 3 columns.

**Format**

This data frame contains the following columns:

- `infect`  Infection time for AIDS, years
- `induct`  Induction time for AIDS, years
- `adult`   Indicator of adult (1=adult, 0=child)

**Source**


**Examples**

```r
data(aids)
```
**alloauto**  
*data from Section 1.9*

**Description**

The alloauto data frame has 90 rows and 5 columns.

**Format**

This data frame contains the following columns:

- **time**  Time to death or relapse, months
- **type**  Type of transplant (1=allogeneic, 2=autologous)
- **delta** Leukemia-free survival indicator (0=alive without relapse, 1=dead or relapse)

**Source**


**Examples**

```r
data(alloauto)
```

---

**allograft**  
*data from Exercise 13.1, p418*

**Description**

The allograft data frame has 34 rows and 4 columns.

**Format**

This data frame contains the following columns:

- **patient**  Patient
- **time**  Time to graft rejection, days
- **rejection**  Indicator of graft rejection (1=yes, 0=no)
- **match**  Good HLA skin match (1=yes, 0=no)

**Source**


**Examples**

```r
data(allograft)
```
azt  

*data from Exercise 4.7, p122*

**Description**

The azt data frame has 45 rows and 4 columns.

**Format**

This data frame contains the following columns:

- **patient**  Patient number
- **ageentry**  Age at entry into AZT study, months
- **age**  Age at death or censoring time, months
- **death**  Death indicator (1=dead, 0=alive)

**Source**


**Examples**

```r
data(azt)
```

baboon  

*data from Exercise 5.8, p147*

**Description**

The baboon data frame has 25 rows and 2 columns.

**Format**

This data frame contains the following columns:

- **date**  Date (day/month/year)
- **time**  Descent time (military time)
- **observed**  Indicator of observed or not (1=observed, 0=not observed)

**Source**


**Examples**

```r
data(baboon)
```
**bcdeter**

---

**data from Section 1.18**

**Description**

The bcdeter data frame has 92 rows and 3 columns.

**Format**

This data frame contains the following columns:

- **lower** Lower limit of interval, months
- **upper** Upper limit of interval, months
- **treat** Treatment regimen (1=radiotherapy only, 2=radiotherapy + chemotherapy)

**Source**


**Examples**

```r
data(bcdeter)
```

---

**bfeed**

---

**data from Section 1.14**

**Description**

The bfeed data frame has 927 rows and 10 columns.

**Format**

This data frame contains the following columns:

- **duration** Duration of breast feeding, weeks
- **delta** Indicator of completed breast feeding (1=yes, 0=no)
- **race** Race of mother (1=white, 2=black, 3=other)
- **poverty** Mother in poverty (1=yes, 0=no)
- **smoke** Mother smoked at birth of child (1=yes, 0=no)
- **alcohol** Mother used alcohol at birth of child (1=yes, 0=no)
- **agemth** Age of mother at birth of child
- **ybirth** Year of birth
- **yschool** Education level of mother (years of school)
- **pc3mth** Prenatal care after 3rd month (1=yes, 0=no)
Source


Examples

data(bfeed)

bmt  

*data from Section 1.3*

Description

The bmt data frame has 137 rows and 22 columns.

Format

This data frame contains the following columns:

- **group**: Disease Group 1-ALL, 2-AML Low Risk, 3-AML High Risk
- **t1**: Time To Death Or On Study Time
- **t2**: Disease Free Survival Time (Time To Relapse, Death Or End Of Study)
- **d1**: Death Indicator 1-Dead 0-Alive
- **d2**: Relapse Indicator 1-Relapsed, 0-Disease Free
- **d3**: Disease Free Survival Indicator 1-Dead Or Relapsed, 0-Alive Disease Free
- **ta**: Time To Acute Graft-Versus-Host Disease
- **da**: Acute GVHD Indicator 1-Developed Acute GVHD 0-Never Developed Acute GVHD
- **tc**: Time To Chronic Graft-Versus-Host Disease
- **dc**: Chronic GVHD Indicator 1-Developed Chronic GVHD 0-Never Developed Chronic GVHD
- **tp**: Time To Chronic Graft-Versus-Host Disease
- **dp**: Platelet Recovery Indicator 1-Platelets Returned To Normal, 0-Platelets Never Returned to Normal
- **z1**: Patient Age In Years
- **z2**: Donor Age In Years
- **z3**: Patient Sex: 1-Male, 0-Female
- **z4**: Donor Sex: 1-Male, 0-Female
- **z5**: Patient CMV Status: 1-CMV Positive, 0-CMV Negative
- **z6**: Donor CMV Status: 1-CMV Positive, 0-CMV Negative
- **z7**: Waiting Time to Transplant In Days
- **z8**: FAB: 1-FAB Grade 4 Or 5 and AML, 0-Otherwise
- **z9**: Hospital: 1-The Ohio State University, 2-Alferd, 3-St. Vincent, 4-Hahnemann
- **z10**: MTX Used as a Graft-Versus-Host-Prophylactic: 1-Yes 0-No
bnct

Source

Examples
data(bmt)

bnct  
*data from Exercise 7.7, p223*

Description
The bnct data frame has 34 rows and 3 columns.

Format
This data frame contains the following columns:

- **trt** Treatment (1=untreated, 2=radiated, 3=radiated + BPA)
- **time** Death time or on-study time, days
- **death** Death indicator (1=dead, 0=alive)

Source

Examples
data(bnct)

btrial  
*data from Section 1.5*

Description
The btrial data frame has 45 rows and 3 columns.

Format
This data frame contains the following columns:

- **time** Time to death or on-study time, months
- **death** Death indicator (0=alive, 1=dead)
- **im** Immunohistochemical response (1=negative, 2=positive)
Source


Examples

data(btrial)

data from Section 1.6

Description

The burn data frame has 154 rows and 17 columns.

Format

This data frame contains the following columns:

- **Obs**: Observation number
- **Z1**: Treatment: 0=routine bathing 1=Body cleansing
- **Z2**: Gender (0=male 1=female)
- **Z3**: Race: 0=nonwhite 1=white
- **Z4**: Percentage of total surface area burned
- **Z5**: Burn site indicator: head 1=yes, 0=no
- **Z6**: Burn site indicator: buttock 1=yes, 0=no
- **Z7**: Burn site indicator: trunk 1=yes, 0=no
- **Z8**: Burn site indicator: upper leg 1=yes, 0=no
- **Z9**: Burn site indicator: lower leg 1=yes, 0=no
- **Z10**: Burn site indicator: respiratory tract 1=yes, 0=no
- **Z11**: Type of burn: 1=chemical, 2=scald, 3=electric, 4=flame
- **T1**: Time to excision or on study time
- **D1**: Excision indicator: 1=yes 0=no
- **T2**: Time to prophylactic antibiotic treatment or on study time
- **D2**: Prophylactic antibiotic treatment: 1=yes 0=no
- **T3**: Time to straphylocous aureaus infection or on study time
- **D3**: Straphylocous aureaus infection: 1=yes 0=no

Source


Examples

data(burn)
channing

**data from Section 1.16**

**Description**

The channing data frame has 462 rows and 6 columns.

**Format**

This data frame contains the following columns:

- **obs** Observation number
- **death** Death status (1=dead, 0=alive)
- **ageentry** Age of entry into retirement home, months
- **age** Age of death or left retirement home, months
- **time** Difference between the above two ages, months
- **gender** Gender (1=male, 2=female)

**Source**


**Examples**

data(channing)

drug6mp

**data from Section 1.2**

**Description**

The drug6mp data frame has 21 rows and 5 columns.

**Format**

This data frame contains the following columns:

- **pair** pair number
- **remstat** Remission status at randomization (1=partial, 2=complete)
- **t1** Time to relapse for placebo patients, months
- **t2** Time to relapse for 6-MP patients, months
- **relapse** Relapse indicator (0=censored, 1=relapse) for 6-MP patients
Source

Examples

```r
data(drug6mp)
```

---

**drughiv**  
*data from Exercise 7.6, p222*

Description
The drughiv data frame has 34 rows and 3 columns.

Format
This data frame contains the following columns:

- **drug** Drug combination (1=A TZ + zalcitabine, 2=A TZ + zalcitabine + saquinavir)
- **time** Time after drug administration to CD4 count at a specified level, days
- **delta** Indicator of CD4 count reaching specified level (1=yes, 0=no)

Source

Examples

```r
data(drughiv)
```

---

**hodg**  
*data from Section 1.10*

Description
The hodg data frame has 43 rows and 6 columns.
**kidney**

**Format**

This data frame contains the following columns:

- **gtype**: Graft type (1=allogenic, 2=autologous)
- **dtype**: Disease type (1=Non Hodgkin lymphoma, 2=Hodgkins disease)
- **time**: Time to death or relapse, days
- **delta**: Death/relapse indicator (0=alive, 1=dead)
- **score**: Karnofsky score
- **wtime**: Waiting time to transplant in months

**Source**


**Examples**

```r
data(hodg)
```

---

**kidney**

*data from Section 1.4*

**Description**

The kidney data frame has 119 rows and 3 columns.

**Format**

This data frame contains the following columns:

- **time**: Time to infection, months
- **delta**: Infection indicator (0=no, 1=yes)
- **type**: Catheter placement (1=surgically, 2=percutaneously)

**Source**


**Examples**

```r
data(kidney)
```
Data on 38 individuals using a kidney dialysis machine

Data on 38 individuals using a kidney dialysis machine See Problem 13.5.2

Usage
data(kidrecurr)

Format
A data frame with 38 observations on the following 10 variables.

patient Patient number
time1 Time one of recurrence of infection, days
infect1 Indicator infection one (1=yes, 0=no)
time2 Time two of recurrence of infection, days
infect2 Indicator infection two (1=yes, 0=no)
age Patient’s age
gender Patient’s gender
gn Disease type GN (1=yes, 0=no)
an Disease type AN (1=yes, 0=no)
pkd Disease type PKD (1=yes, 0=no)

Source

Examples
data(kidrecurr)
Kidtran

*Description*

The `kidtran` data frame has 863 rows and 6 columns.

*Format*

This data frame contains the following columns:

- **obs**: Observation number
- **time**: Time to death or on-study time
- **delta**: Death indicator (0=alive, 1=dead)
- **gender**: 1=male, 2=female
- **race**: 1=white, 2=black
- **age**: Age in years

*Source*


*Examples*

```r
data(kidtran)
```

Larynx

*Description*

The `larynx` data frame has 90 rows and 5 columns.

*Format*

This data frame contains the following columns:

- **stage**: Stage of disease (1=stage 1, 2=stage2, 3=stage 3, 4=stage 4)
- **time**: Time to death or on-study time, months
- **age**: Age at diagnosis of larynx cancer
- **diagyr**: Year of diagnosis of larynx cancer
- **delta**: Death indicator (0=alive, 1=dead)
lifetab

Source


Examples

data(larynx)

lifetab

*Create cohort life table*

Description

Create cohort life table.

Usage

lifetab(tis, ninit, nlost, nevent)

Arguments

tis          a vector of end points of time intervals, whose length is 1 greater than nlost and nevent.
ninit       the number of subjects initially entering the study.
nlost        a vector of the number of individuals lost follow or withdrawn alive for whatever reason.
nevent       a vector of the number of individuals who experienced the event

Value

A data.frame with the following columns:

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nsubs</td>
<td>the number of subject entering the intervals who have not experienced the event.</td>
</tr>
<tr>
<td>nlost</td>
<td>the number of individuals lost follow or withdrawn alive for whatever reason.</td>
</tr>
<tr>
<td>nrisk</td>
<td>the estimated number of individuals at risk of experiencing the event.</td>
</tr>
<tr>
<td>nevent</td>
<td>the number of individuals who experienced the event.</td>
</tr>
<tr>
<td>surv</td>
<td>the estimated survival function at the start of the intervals.</td>
</tr>
<tr>
<td>pdf</td>
<td>the estimated probability density function at the midpoint of the intervals.</td>
</tr>
<tr>
<td>hazard</td>
<td>the estimated hazard rate at the midpoint of the intervals.</td>
</tr>
<tr>
<td>se.surv</td>
<td>the estimated standard deviation of survival at the beginning of the intervals.</td>
</tr>
<tr>
<td>se.pdf</td>
<td>the estimated standard deviation of the probability density function at the midpoint of the intervals.</td>
</tr>
<tr>
<td>se.hazard</td>
<td>the estimated standard deviation of the hazard function at the midpoint of the intervals.</td>
</tr>
</tbody>
</table>

The row.names are the intervals.
Author(s)

Jun Yan <jyan@stat.uconn.edu>

Examples

tis <- c(0, 2, 3, 5, 7, 11, 17, 25, 37, 53, NA)
nsubs <- c(927, 848, 774, 649, 565, 449, 296, 186, 112, 27)
nlost <- c(2, 3, 6, 9, 7, 5, 3, rep(0, 3))
nevent <- c(77, 71, 119, 75, 109, 148, 107, 74, 85, 27)
lifetab(tis, nsubs[1], nlost, nevent)

data(lung)

Description

The lung data frame has 25 rows and 4 columns.

Format

This data frame contains the following columns:

- **time**  Days to death
- **death** Death indicator (1=dead), complete follow-up on all patients
- **time2** Days to 3/31/80 or death (interim analysis)
- **death2** Death indicator as of 3/31/80 (1=dead, 0=alive)

Source


Examples

data(lung)
The pneumon data frame has 3470 rows and 15 columns.

This data frame contains the following columns:

- **chldage**: Age child had pneumonia, months
- **hospital**: Indicator for hospitalization for pneumonia (1=yes, 0=no)
- **mthage**: Age of the mother, years
- **urban**: Urban environment for mother (1=yes, 0=no)
- **alcohol**: Alcohol use by mother during pregnancy (1=yes, 0=no)
- **smoke**: Cigarette use by mother during pregnancy (1=yes, 0=no)
- **region**: Region of the country (1=northeast, 2=north central, 3=south, 4=west)
- **poverty**: Mother at poverty level (1=yes, 0=no)
- **bweight**: Normal birthweight (>5.5 lbs.) (1=yes, 0=no)
- **race**: Race of the mother (1=white, 2=black, 3=other)
- **education**: Education of the mother, years of school
- **nsibs**: Number of siblings of the child
- **wmonth**: Month the child was weaned
- **sfmonth**: Month the child on solid food
- **agepn**: Age child in the hospital for pneumonia, months


Examples

```r
data(pneumon)
```
psych

data from Section 1.15

Description
The psych data frame has 927 rows and 10 columns.

Format
This data frame contains the following columns:

- **sex**: Patient sex (1=male, 2=female)
- **age**: Patient age
- **time**: Time to death or on-study time
- **death**: Death indicator (0=alive, 1=dead)

Source

Examples
```
data(psych)
```

rats

data from Exercise 7.13, p225

Description
The rats data frame has 50 rows and 4 columns.

Format
This data frame contains the following columns:

- **time**: Time to tumor development
- **tumor**: Indicator of tumor development (1=yes, 0=no)
- **trt**: Treatment (1=treated with drug, 0=given placebo)
- **litter**: Litter

Source
Examples

data(rats)

---

std

data from Section 1.12

---

Description

The std data frame has 877 rows and 3 columns.

Format

This data frame contains the following columns:

- **obs** Observation number
- **race** Race (W=white, B=black)
- **marital** Marital status (D=divorced / separated, M=married, S=single)
- **age** AGE
- **yschool** Years of schooling
- **iinfct** Initial infection (1= gonorrhea, 2=chlamydia, 3=both)
- **npartner** Number of partners
- **os12m** Oral sex within 12 months (1=yes, 0=no)
- **os30d** Oral sex within 30 days (1=yes, 0=no)
- **rs12m** Rectal sex within 12 months (1=yes, 0=no)
- **rs30d** Rectal sex within 30 days (1=yes, 0=no)
- **abdpain** Presence of abdominal pain (1=yes, 0=no)
- **discharge** Sign of discharge (1=yes, 0=no)
- **dysuria** Sign of dysuria (1=yes, 0=no)
- **condom** Condom use (1=always, 2=sometime, 3=never)
- **itch** Sign of itch (1=yes, 0=no)
- **lesion** Sign of lesion (1=yes, 0=no)
- **rash** Sign of rash (1=yes, 0=no)
- **lymph** Sign of lymph (1=yes, 0=no)
- **vagina** Involvement vagina at exam (1=yes, 0=no)
- **dchexam** Discharge at exam (1=yes, 0=no)
- **abnode** Abnormal node at exam (1=yes, 0=no)
- **rinfct** Reinfection (1=yes, 0=no)
- **time** Time to reinfection
**Source**


**Examples**

data(std)

tongue

data(stddiag)

data(tongue)

**Description**

The *stddiag* data frame has 25 rows and 2 columns.

**Format**

This data frame contains the following columns:

- **encounter** Months from 1/93 to encounter
- **diagnosed** Months until STD diagnosed in the clinic

**Source**


**Examples**

data(stddiag)

tongue

data(tongue)

**Description**

The *tongue* data frame has 80 rows and 3 columns.

**Format**

This data frame contains the following columns:

- **type** Tumor DNA profile (1=Aneuploid Tumor, 2=Diploid Tumor)
- **time** Time to death or on-study time, weeks
- **delta** Death indicator (0=alive, 1=dead)
20

Source


Examples

data(tongue)

twins data from Exercise 7.14, p225

Description

The twins data frame has 24 rows and 3 columns.

Format

This data frame contains the following columns:

- **id**: Twin number
- **age**: Age of twin’s death from CHD, months
- **death**: Death (male twin) from CHD indicator (1=dead from CHD, 0=alive or other cause of death)
- **gender**: 1=male, 2=female

Source


Examples

data(twins)
# Index

*Topic **datasets**
- aids, 2
- alloauto, 3
- allograft, 3
- azt, 4
- baboon, 4
- bcdeter, 5
- bfeed, 5
- bmt, 6
- bnct, 7
- btrial, 7
- burn, 8
- channing, 9
- drug6mp, 9
- drughiv, 10
- hodg, 10
- kidney, 11
- kidrecurr, 12
- kidtran, 13
- larynx, 13
- lung, 15
- pneumon, 16
- psych, 17
- rats, 17
- std, 18
- stddiag, 19
- tongue, 19
- twins, 20

*Topic **manip**
- lifetab, 14

aids, 2
alloauto, 3
allograft, 3
azt, 4
baboon, 4
bcdeter, 5
bfeed, 5
bmt, 6
bnct, 7
btrial, 7
burn, 8
channing, 9
drug6mp, 9
drughiv, 10
hodg, 10
kidney, 11
kidrecurr, 12
kidtran, 13
larynx, 13
lung, 15
pneumon, 16
psych, 17
rats, 17
std, 18
stddiag, 19
tongue, 19
twins, 20