Package 'asaur'

April 12, 2016

Type Package

Title Data Sets for ``Applied Survival Analysis Using R''"

Version 0.50

LazyData true

Date 2016-04-10

Author Dirk F. Moore

Description Data sets are referred to in the text ``Applied Survival Analysis Using R" by Dirk F. Moore, Springer, 2016, ISBN: 978-3-319-31243-9, <DOI:10.1007/978-3-319-31245-3>.

Maintainer Dirk F. Moore <dirkfmoore@gmail.com>

License CC0

Repository CRAN

NeedsCompilation no

Date/Publication 2016-04-12 06:23:05

R topics documented:

ashkenazi *ashkenazi*

Description

This is a random subset of data from the Struewing et al. (1997) study of Ashkenazi jews and breast cancer. The subset consists of pairs of first-degree female relatives who are also first degree relatives of a proband.

Usage

```
data("ashkenazi")
```

Format

A data frame with 3920 observations on the following 4 variables.

`famID` family ID indicator

`brcancer` 1 if subject had breast cancer, 0 if not

`age` Age at onset of breast cancer, or current age if no breast cancer

`mutant` 1 if first degree relative proband was a BRCA mutation carrier, 0 if not

References

Moore DF, Chatterjee N, Pee D, and Gail MH (2001) Pseudo-likelihood estimates of the cumulative risk of an autosomal dominant disease from a kin-cohort study. Genetic Epidemiology 20, 210-227.)

Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC, and Tucker MA (1997) The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among ashkenazi jews. New England Journal of Medicine 336, 1401-1408.)

Examples

```
data(ashkenazi)
```

ChanningHouse *Channing House Data*

Description

The ChanningHouse data frame has 457 rows and 5 columns. This is 5 fewer than the parent channing data frame in the boot package. These 5 were removed because the exit time was not smaller than the entry time.

Channing House is a retirement centre in Palo Alto, California. These data were collected between the opening of the house in 1964 until July 1, 1975. In that time 97 men and 365 women passed through the centre. For each of these, their age on entry and also on leaving or death was recorded. A large number of the observations were censored mainly due to the resident being alive on July 1, 1975 when the data was collected. Over the time of the study 130 women and 46 men died at Channing House. Differences between the survival of the sexes, taking age into account, was one of the primary concerns of this study.

Usage

```
data("ChanningHouse")
```

Format

A data frame with 457 observations on the following 5 variables.

`sex` a factor for the sex of each resident with levels `Female` `Male`

`entry` The residents age (in months) on entry to the center)

`exit` The age (in months) of the resident on death, leaving the center or July 1, 1975, whichever event occurred first.)

`time` The length of time (in months) that the resident spent at Channing House. (`time=exit-entry`)))

`cens` The indicator of reight censoring. 1 indicates that the resident died at Channing House, 0 indicates that they left the house prior to July 1, 1975 or that they were still alive and living in the center at that date.

Source

The current data were derived from the "channing" data frame in the "boot" package. The original source for the data was

Hyde, J. (1980) Testing survival with incomplete observations. Biostatistics Casebook. R.G. Miller, B. Efron, B.W. Brown and L.E. Moses (editors), 31-46. John Wiley.

References

Davison, A.C. and Hinkley, D.V. (1997) Bootstrap Methods and Their Application. Cambridge University Press.

Canty, A. and Ripley, B. (2015) boot package.

Examples

```
data(ChanningHouse)
```

gastricXelox *gasticXelox*

Description

Data from a Phase II clinical trial of Xeloda and exaliplatin given before surgery to advanced gastric cancer patients with para-aortic lymph node metastasis.

Usage

```
data("gastricXelox")
```

Format

A data frame with 48 observations on the following 2 variables.

`timeWeeks` survival time in weeks

`delta` 1 for death, 0 for censored

Details

The data were extracted from the Kaplan-Meier survival plot.

References

Wang Y, Yu Y-Y, Li W, Feng Y, Hou J, Ji Y, Sun Y-H, Shen K-T, Shen Z-B, Qin X-Y, and Liu T-S. (2014) A phase II trial of xeloda and oxaliplatin (XELOX) neo-adjuvant chemotherapy followed by surgery for advanced gastric cancer patients with para-aortic lymph node metastasis. Cancer Chemotherapy and Pharmacology 73(6), 1155-1161.))

Examples

```
data(gastricXelox)
```

hepatoCellular *hepatoCellular*

Description

Overall and recurrence-free survival of patients with hepatocellular carcinoma.

Usage

```
data("hepatoCellular")
```

Format

A data frame with 227 observations on 48 clinical and biomarker variables

`Number` Patient ID number

`Age` a numeric vector

`Gender` a numeric vector

`HBsAg` a numeric vector

`Cirrhosis` a numeric vector

`ALT` a numeric vector

`AST` a numeric vector

`AFP` a numeric vector

`Tumorsize` a numeric vector

`Tumordifferentiation` a numeric vector

`Vascularinvasion` a numeric vector

`Tumormultiplicity` a numeric vector

`Capsulation` a numeric vector

`TNM` a numeric vector

`BCLC` a numeric vector

`OS` Overall survival

`Death` 1 denotes death, 0 censored

`RFS` Recurrence-free survival

`Recurrence` 1 denotes recurrence, 0 censored

`CXCL17T` a numeric vector

`CXCL17P` a numeric vector

`CXCL17N` a numeric vector

`CD4T` a numeric vector

`CD4N` a numeric vector

`CD8T` a numeric vector

`CD8N` a numeric vector

`CD20T` a numeric vector

`CD20N` a numeric vector

`CD57T` a numeric vector

`CD57N` a numeric vector

`CD15T` a numeric vector

`CD15N` a numeric vector

`CD68T` a numeric vector

`CD68N` a numeric vector

`CD4NR` a numeric vector

CD8NR a numeric vector

CD20NR a numeric vector

CD57NR a numeric vector

CD15NR a numeric vector

CD68NR a numeric vector

CD4TR a numeric vector

CD8TR a numeric vector

CD20TR a numeric vector

CD57TR a numeric vector

CD15TR a numeric vector

CD68TR a numeric vector

Ki67 a numeric vector

CD34 a numeric vector

References

Li L, Yan J, Xu J, Liu C-Q, Zhen Z-J, Chen H-W, Ji Y, Wu Z-P, Hu J-Y, Zheng L, Lau WY (2014) Cxcl17 expression predicts poor prognosis and correlates with adverse immune infiltration in hepatocellular carcinoma. Plos One 9 (10) e110064.

Li L, Yan J, Xu J, Liu C-Q, Zhen Z-J, Chen H-W, Ji Y, Wu Z-P, Hu J-Y, Zheng L, Lau WY (2014) Cxcl17 expression predicts poor prognosis and correlates with adverse immune infiltration in hepatocellular carcinoma. Dryad Digital Repository datadryad.org.

Examples

```
data(hepatoCellular)
```

---

pancreatic *pancreatic*

Description

Data from a Phase II clinical trial of patients with locally advanced or metastatic pancreatic cancer.

Usage

```
data("pancreatic")
```

Format

A data frame with 41 observations on the following 4 variables.

`stage` a factor with levels `LA` (locally advanced) or `M` (metastatic)

`onstudy` date of enrollment into the clinical trial, in month/day/year format

`progression` date of progression, in month/day/year format

`death` date of death, in month/day/year format

Details

Since all patients in this study have known death dates, there is no censoring.

References

Moss RA, Moore D, Mulcahy MF, Nahum K, Saraiya B, Eddy S, Kleber M, and Poplin EA (2012) A multi-institutional phase 2 study of imatinib mesylate and gemcitabine for first-line treatment of advanced pancreatic cancer. Gastrointestinal Cancer Research 5, 77 - 83.

Examples

```
data(pancreatic)
```

`pancreatic2` *pancreatic2*

Description

This is the same data as in 'pancreatic', with overall and progression-free survival calculated. Dates have been removed.

Usage

```
data("pancreatic2")
```

Format

A data frame with 41 observations on the following 4 variables.

`pfs` Progression-free survival: Time from entry until disease progression. If no progression was observed, before death, the time to death is used.

`os` Overall survival: Time from entry until death

`status` This censoring indicator is 1 for all patients, since all patients died.

`stage` a factor with levels `LA` (locally advanced) or `M` (metastatic)

References

Moss RA, Moore D, Mulcahy MF, Nahum K, Saraiya B, Eddy S, Kleber M, and Poplin EA (2012) A multi-institutional phase 2 study of imatinib mesylate and gemcitabine for first-line treatment of advanced pancreatic cancer. Gastrointestinal Cancer Research 5, 77 - 83.

Examples

```
data(pancreatic2)
```

`pharmacoSmoking` *pharmacoSmoking*

Description

Randomized trial of triple therapy vs. patch for smoking cessation.

Usage

```
data("pharmacoSmoking")
```

Format

A data frame with 125 observations on the following 14 variables.

`id` patient ID number

`ttr` Time in days until relapse

`relapse` Indicator of relapse (return to smoking)

`grp` Randomly assigned treatment group with levels `combination` or `patchOnly`

`age` Age in years at time of randomization

`gender` `Female` or `Male`

`race` `black`, `hispanic`, `white`, or `other`

`employment` `ft` (full-time), `pt` (part-time), or `other`

`yearsSmoking` Number of years the patient had been a smoker

`levelSmoking` `heavy` or `light`

`ageGroup2` Age group with levels `21-49` or `50+`

`ageGroup4` Age group with levels `21-34`, `35-49`, `50-64`, or `65+`

`priorAttempts` The number of prior attempts to quit smoking

`longestNoSmoke` The longest period of time, in days, that the patient has previously gone without smoking

Source

This data is from a clinical trial described in Steinberg et al. (2009)

References

Steinberg, M.B. Greenhaus, S. Schmelzer, A.C. Bover, M.T., Foulds, J., Hoover, D.R., and Carson, J.L. (2009) Triple-combination pharmacotherapy for medically ill smokers: A randomized trial. Annals of Internal Medicine 150, 447-454.

Examples

```
data(pharmacoSmoking)
```

`prostateSurvival` *prostateSurvival*

Description

This data set contains survival times for two competing causes: time from prostate cancer diagnosis to death from prostate cancer, and time from prostate cancer diagnosis to death from other causes. The data set also contains information on several risk factors. The data in this data set are simulated from detailed competing risk survival curves and counts of numbers of patients per group presented in Lu-Yao et al. (2009). Thus, the simulated data presented here contain many of the characteristics of the original SEER-Medicare prostate cancer data used in Lu-Yao et al. (2009).

Usage

```
data("prostateSurvival")
```

Format

A data frame with 14294 observations on the following 5 variables.

`grade` a factor with levels `mode` (moderately differentiated) and `poor` (poorly differentiated)

`stage` a factor with levels `T1ab` (Stage T1, clinically diagnoseed), `T1c` (Stage T1, diagnosed via a PSA test), and `T2` (Stage T2)

`ageGroup` a factor with levels `66-69` `70-74` `75-79` `80+`

`survTime` time from diagnosis to death or last date known alive

`status` a censoring variable, `0`, (censored), `1` (death from prostate cancer), and `2` (death from other causes)

Source

Lu-Yao, GL, Albertsen PC, Moore DF, Shih W, Lin Y, DiPaola RS, Barry MJ, Zietman A, O'Leary M, Walker-Corkery E, Yao S-L (2009) Outcomes of localized prostate cancer following conservative management. Journal of the American Medical Association 302, 1202 - 1209.)

Examples

```
data(prostateSurvival)
```

Index