

Package ‘parfm’

January 17, 2023

Type Package

Title Parametric Frailty Models

Version 2.7.7

Description Fits Parametric Frailty Models by maximum marginal likelihood.

Possible baseline hazards:

exponential, Weibull, inverse Weibull (Fréchet),
Gompertz, lognormal, log-skew-normal, and loglogistic.

Possible Frailty distributions:

gamma, positive stable, inverse Gaussian and lognormal.

License GPL-2

LazyLoad yes

Depends R (>= 2.10), survival, optimx

Imports msm, graphics, sn

Encoding UTF-8

NeedsCompilation no

Author Federico Rotolo [aut, cre],

Marco Munda [aut],

Andrea Callegaro [ctb]

Maintainer Federico Rotolo <federico.rotolo@sanofi.com>

Repository CRAN

Date/Publication 2023-01-17 22:40:02 UTC

R topics documented:

anova.parfm	2
asthma	3
ci.parfm	5
culling	6
diagnosis	7
ecf	9
insem	11

insemtvc	14
mastitis	15
parfm	17
predict.parfm	21
reconstitution	23
select.parfm	25
tau	26

Index	28
--------------	-----------

anova.parfm	<i>Analysis of Deviance for a parametric frailty model.</i>
-------------	---

Description

Compute an analysis of deviance table for one or more parametric frailty model fits.

Usage

```
## S3 method for class 'parfm'
anova(object, ...)
```

Arguments

object	An object of class parfm
...	Further parfm objects

Details

Specifying a single object gives a sequential analysis of deviance table for that fit. That is, the reductions in the model log-likelihood as each term of the formula is added in turn are given in as the rows of a table, plus the log-likelihoods themselves.

If more than one object is specified, the table has a row for the degrees of freedom and loglikelihood for each model. For all but the first model, the change in degrees of freedom and loglik is also given. (This only make statistical sense if the models are nested.) It is conventional to list the models from smallest to largest, but this is up to the user.

The table contains test statistics (and P values) comparing the reduction in loglik for each row.

Value

An object of class "anova" inheriting from class "data.frame".

Warning

The comparison between two or more models by `anova.parfm()` will only be valid if they are fitted to the same dataset. This may be a problem if there are missing values.

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Munda M, Rotolo F, Legrand C (2012). parfm: Parametric Frailty Models in R. Journal of Statistical Software, 51(11), 1-20. DOI <doi: [10.18637/jss.v051.i11](https://doi.org/10.18637/jss.v051.i11)>

See Also

[parfm](#), [anova](#).

Examples

```
fit <- parfm(formula = Surv(time, status) ~ sex + age, cluster = "id",
            data = kidney, dist = "exponential", frailty = "gamma")
anova(fit)
fit2 <- parfm(formula = Surv(time, status) ~ sex, cluster = "id",
            data = kidney, dist = "exponential", frailty = "gamma")
anova(fit2, fit)
```

asthma

Recurrent asthma attacks in children

Description

Asthma is occurring more and more frequently in very young children (between 6 and 24 months). Therefore, a new application of an existing anti-allergic drug is administered to children who are at higher risk to develop asthma in order to prevent it. A prevention trial is set up with such children randomised to placebo or drug, and the asthma events that developed over time are recorded in a diary. Typically, a patient has more than one asthma event. The different events are thus clustered within a patient and are ordered in time. This ordering can be taken into account in the model. Such data can be presented in different formats, but here, we choose to use the calendar time representation. In the calendar time representation, the time at risk for a particular event is the time from the end of the previous event (asthma attack) to the start of the next event (start of the next asthma attack). In describing recurrent event data, we need a somewhat more complex data structure to keep track of the sequence of events within a patient. A particular patient has different periods at risk during the total observation period which are separated either by an asthmatic event that lasts one or more days or by a period in which the patient was not under observation. The start and end of each such risk period is required, together with the status indicator to denote whether the end of the risk period corresponds to an asthma attack or not.

Usage

```
data(asthma)
```

Format

A dataframe containing 1776 observations.

Patid: Patient's identifier.

Begin: Time of end of the previous asthma attack (in days).

End: Asthma attack or censoring time. (in days)

Status: Censored (0) or observed (1) event time.

Drug: placebo (0) or drug (1).

Fevent: First observation of the patient? 1=yes, 0=no.

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Author(s)

F. Rotolo and M. Munda. Original text and data by L. Duchateau and P. Janssen.

Source

Example 1.9 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```
data(asthma)
head(asthma)
asthma <- asthma[asthma$Fevent==0,]

#####
# Example 2.4: The frailty model with the Weibull baseline for the recurrent #
# asthma data based on marginal likelihood maximisation #
# Duchateau and Janssen (2008, page 56) #
#####
# Calendar time
parfm(Surv(Begin, End, Status) ~ Drug, cluster = "Patid", data = asthma,
      dist = "weibull", frailty = "gamma")

# Gap time
asthma$time <- asthma$End - asthma$Begin
parfm(Surv(time, Status) ~ Drug, cluster = "Patid", data = asthma,
      dist = "weibull", frailty = "gamma")
```

ci.parfm	<i>Confidence Intervals for Hazard Ratios of Covariates of Parametric Frailty Models</i>
----------	--

Description

Computes confidence intervals for hazard ratios ($\exp(\text{coef})$) for objects of class `parfm`.

Usage

```
ci.parfm(x, level=.05, digits=3)
```

Arguments

x	The fitted model, object of class <code>parfm</code> , obtained as result of the <code>parfm()</code> function.
level	The coverage probability of the confidence interval.
digits	The number of significant digits.

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Munda M, Rotolo F, Legrand C (2012). `parfm`: Parametric Frailty Models in R. *Journal of Statistical Software*, 51(11), 1-20. DOI <doi: [10.18637/jss.v051.i11](https://doi.org/10.18637/jss.v051.i11)>

See Also

[parfm](#), [select.parfm](#), [predict.parfm](#)

Examples

```
data("kidney")
kidney$sex <- kidney$sex - 1

pfm <- parfm(Surv(time,status) ~ sex + age, cluster="id",
             data=kidney, dist="exponential", frailty="gamma")

ci.parfm(pfm)
```

culling

Culling of dairy heifer cows

Description

The time to culling is studied in heifers as a function of the somatic cell count (SCC) measured between 5 and 15 days (measurement day) after calving.

High somatic cell count (we use the logarithm of somatic cell count as covariate) might be a surrogate marker for intramammary infections. Heifers which have intramammary infections or which are expected to develop intramammary infections in the future are quite expensive to keep due to the high costs for drugs and the loss in milk production.

Cows are followed up for an entire lactation period (roughly 300–50 days) and, if they are still alive at the end of the lactation period, they are censored at that time. Cows are further clustered within herds and this clustering needs to be taken into account as culling policy and also SCC in early lactation might differ substantially between the herds.

Usage

```
data(culling)
```

Format

A dataframe containing 13836 observations.

Cowid: Cow's identifier.

Time: Time to culling (in days).

Status: Censored (0) or observed (1) event time.

Herd: Herd identifier.

Timeasses: SCC measurement day.

LogSCC: Logarithm of the somatic cell count.

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.7 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```
data(culling)
head(culling)
culling <- culling[culling$Time > 0,]
culling$TimeMonths <- culling$Time * 12 / 365.25

coxmod <- parfm(Surv(TimeMonths, Status) ~ LogSCC, data = culling,
               dist = "exponential", frailty = "none")
coxmod

pfmod <- parfm(Surv(TimeMonths, Status) ~ LogSCC, data = culling,
               cluster = "Herd", dist = "exponential", frailty = "gamma")
pfmod
```

diagnosis

Diagnosis of fracture healing

Description

Medical imaging has become an important tool in the veterinary hospital to assess whether and when a fracture has healed. The standard technique in dogs is based on radiography (RX). Newer techniques based on ultrasound (US) are cheaper and do not require radioprotection.

To investigate the performance of US for this purpose and to compare it to RX, Risselada et al. (2006) set up a trial in which fracture healing is evaluated by both US and RX. In total, 106 dogs, treated in the veterinary university hospital of Ghent, are included in the trial and evaluated for time to fracture healing with the two techniques. Only 7 dogs are censored for time to fracture healing evaluated by RX; no censoring occurs for time to fracture healing evaluated by US. The censoring is due to the fact that dog owners do not show up anymore.

Usage

```
data(diagnosis)
```

Format

A dataframe containing 212 observations.

Dogid: Dog's identifier.

Time: Time to diagnosis (in days).

Status: Censored (0) or observed (1) event time.

Method: Diagnostic technique: either RX, radiography, or US, ultrasound.

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.2 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Risselada M, van Bree H, Kramer M, Chiers K, Duchateau L, Verleyen P (2006). Evaluation of nonunion fractures in dogs by use of Bmode ultrasonography, power Doppler ultrasonography, radiography, and histologic examination. *Am. J. Vet. Res.* 67, 1354–1361.

Examples

```

data(diagnosis)
head(diagnosis)

diagnosis$TimeMonths <- diagnosis$Time * 12 / 365.25

#####
# Example 3.6: Shared gamma frailty models [...] for time to diagnosis      #
# of being healed                                                            #
# Duchateau and Janssen (2008, page 101)                                     #
#####
WeiGam <- parfm(Surv(TimeMonths, Status) ~ Method,
                cluster = "Dogid", data = diagnosis,
                dist = "weibull", frailty = "gamma")
WeiGam

curve(WeiGam["lambda", 1] * WeiGam["rho", 1] * x ^ (WeiGam["rho", 1] - 1),
      from = 0, to = 4, ylim = c(0, 25),
      ylab = "Hazard function", xlab = "Time (months)")
curve(exp(WeiGam["Method", 1]) *
      WeiGam["lambda", 1] * WeiGam["rho", 1] * x ^ (WeiGam["rho", 1] - 1),
      add = TRUE, lty = 2)
legend("topleft", lty = 1:2, legend = c("US", "RX"))

#####
# Example 4.8: Inverse Gaussian frailty models [...] for time to diagnosis  #
# of being healed                                                            #
# Duchateau and Janssen (2008, page 160)                                     #
#####
WeiIG <- parfm(Surv(TimeMonths, Status) ~ Method,
                cluster = "Dogid", data = diagnosis,
                dist = "weibull", frailty = "ingau")
WeiIG

curve(WeiIG["lambda", 1] * WeiIG["rho", 1] * x ^ (WeiIG["rho", 1] - 1),
      from = 0, to = 4, ylim = c(0, 25),
      ylab = "Hazard function", xlab = "Time (months)")
curve(exp(WeiIG["Method", 1]) *
      WeiIG["lambda", 1] * WeiIG["rho", 1] * x ^ (WeiIG["rho", 1] - 1),

```



```

      add = TRUE, lty = 2)
legend("topleft", lty = 1:2, legend = c("US", "RX"))

#####
# Example 4.11: Positive Stable frailty models [...] for time to diagnosis      #
# of being healed                                                            #
# Duchateau and Janssen (2008, page 172)                                     #
#####
WeiPS <- parfm(Surv(TimeMonths, Status) ~ Method,
               cluster = "Dogid", data = diagnosis,
               dist = "weibull", frailty = "possta")

WeiPS

curve(WeiPS["lambda", 1] * WeiPS["rho", 1] * x ^ (WeiPS["rho", 1] - 1),
      from = 0, to = 4, ylim = c(0, 25),
      ylab = "Hazard function", xlab = "Time (months)")
curve(exp(WeiPS["Method", 1]) *
      WeiPS["lambda", 1] * WeiPS["rho", 1] * x ^ (WeiPS["rho", 1] - 1),
      add = TRUE, lty = 2)
legend("topleft", lty = 1:2, legend = c("US", "RX"))

```

ecf

East Coast Fever transmission dynamics

Description

Theileriosis or East Coast Fever (ECF) is a major cattle disease in East and Southern Africa. The disease is caused by *Theileria parva* which is transmitted by the ticks *Rhipicephalus appendiculatus* and the closely related *Rhipicephalus zambeziensis*.

In order to study the transmission of the disease, cows were followed up from birth until time of first ECF contact in Nteme, a region in Southern Zambia. On weekly basis, blood is collected and tested for the presence of antibodies to *Theileria parva* using the Indirect Fluorescent Antibody test. After three consecutive positive IFA test results an animal is considered seroconverted. The time to first ECF contact is defined as the timespan from birth to one month before the first of the three consecutive positive test results (i.e., we set the date of contact with *Theileria parva* one month before the time of the first positive test result). Animals are followed up until one year after birth; if they do not seroconvert by that time, their time of seroconversion is right-censored. We also consider the binary covariate breed.

Usage

```
data(ecf)
```

Format

A dataframe containing 212 observations.

Cowid: Cow's identifier.

Time: Time to ECF contact (in days).

Status: Censored (0) or observed (1) event time.

Breed: The cow's breed.

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.1 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```
data(ecf)
head(ecf)

#####
#Example 3.8: Population and conditional hazard for time to ECF contact      #
#Duchateau and Janssen (2008, page 113)                                    #
#####
pfm1 <- parfm(Surv(Time, Status) ~ 1, cluster = "Cowid", data = ecf,
             dist = "weibull", frailty = "gamma")
pfm2 <- parfm(Surv(Time, Status) ~ Breed, cluster = "Cowid", data = ecf,
             dist = "weibull", frailty = "gamma")
curve(pfm1["lambda", 1] * pfm1["rho", 1] * x ^ (pfm1["rho", 1] - 1),
      from = 0, to = 400, ylim = c(0, .15),
      ylab = "Hazard function", xlab = "Time (days)")
curve(qgamma(.75, shape = 1 / pfm1["theta", 1],
            scale = pfm1["theta", 1] * pfm1["lambda", 1] * pfm1["rho", 1] *
            x ^ (pfm1["rho", 1] - 1),
      add = TRUE, lty = 2)
curve(qgamma(.25, shape=1 / pfm1["theta", 1],
            scale = pfm1["theta", 1] * pfm1["lambda", 1] * pfm1["rho", 1] *
            x ^ (pfm1["rho", 1] - 1),
      add = TRUE, lty = 3)
curve(pfm1["lambda", 1] * pfm1["rho", 1] *
      x ^ (pfm1["rho", 1] - 1) / (
        1 + pfm1["theta", 1] * pfm1["lambda", 1] * x ^ (pfm1["rho", 1])
      ),
      add = TRUE, lwd = 2)
legend("top", lwd = c(1, 1, 1, 2), lty = c(1, 2, 3, 1), ncol = 2,
      legend = c("Mean frailty", "Q75 frailty", "Q25 frailty", "Population"))
```

insem	<i>Time to first insemination in dairy heifer cows without time varying covariates</i>
-------	--

Description

In a dairy farm, the calving interval (the time between two calvings) should be optimally between 12 and 13 months. One of the main factors determining the length of the calving interval is the time from parturition to the time of first insemination.

The objective of this study is to look for cow factors that might predict the time to first insemination, so that actions can be taken based on these predictors. As no inseminations take place in the first 29 days after calving, we subtract 29 days (and not 30 days as the first event would then have first insemination time zero) from the time to first insemination since at risk time starts only then. Cows which are culled without being inseminated are censored at their culling time. Furthermore, cows that are not yet inseminated 300 days after calving are censored at that time.

The time to first insemination is studied in dairy cows as a function of covariates that are fixed over time. An example of such a covariate is the parity of the cow, corresponding to the number of calvings the cow has had already. As we observe only one lactation period for each cow in the study, it is indeed a constant cow characteristic within the time framework of the study.

We dichotomise parity into primiparous cows or heifers (only one calving (`heifer=1`)) and multiparous cows (more than one calving (`heifer=0`)). Other covariates that are used in the analysis are the different milk constituents such as protein and ureum concentration at parturition.

Usage

```
data(insem)
```

Format

A dataframe containing 10513 observations.

Cowid: Cow's identifier.

Time: Time to first insemination (in days).

Status: Censored (0) or observed (1) event time.

Herd: The herd to which the cow belongs.

Urem: Milk ureum concentration (%) at the start of the lactation period.

Protein: Protein concentration (%) at the start of the lactation period.

Parity: The number of calvings.

Heifer: Multiparous cow (0) or primiparous cow (1).

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.8 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```

data(insem)
head(insem)

insem$TimeMonths <- insem$Time * 12 / 365.25

#####
#Example 2.1: The parametric proportional hazards frailty model for the time #
#to first insemination based on marginal likelihood maximisation          #
#Duchateau and Janssen (2008, page 50)                                   #
#####
pfm <- parfm(Surv(TimeMonths, Status) ~ Heifer, cluster = "Herd", data = insem,
             dist = "weibull", frailty = "gamma")

pfm

par(mfrow = c(2, 2))

### - Hazard functions - ###
# multiparous cows
curve((365.25 / 12) ^ (-pfm["rho", 1]) *
      pfm["lambda", 1] * pfm["rho", 1] * x ^ (pfm["rho", 1] - 1),
      from = 0, to = 400, ylim = c(0, .14),
      main = "Multiparous cows",
      ylab = "Hazard function", xlab = "Time to first insemination (days)")
curve(qgamma(.95, shape = 1 / pfm["theta", 1],
             scale = pfm["theta", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
      pfm["lambda", 1] * pfm["rho", 1] * x ^ (pfm["rho", 1] - 1),
      add = TRUE, lty = 4)
curve(qgamma(.05, shape = 1 / pfm["theta", 1],
             scale = pfm["theta", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
      pfm["lambda", 1] * pfm["rho", 1] * x ^ (pfm["rho", 1] - 1),
      add = TRUE, lty = 4)

# primiparous cows
curve(exp(pfm["Heifer", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
      pfm["lambda", 1] * pfm["rho", 1] * x ^ (pfm["rho", 1] - 1),
      from = 0, to = 400, ylim = c(0, .14),
      main = "Primiparous cows",
      ylab = "Hazard function", xlab = "Time to first insemination (days)")
curve(qgamma(.95, shape = 1 / pfm["theta", 1],
             scale = pfm["theta", 1]) * exp(pfm["Heifer", 1]) *
      (365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1] * pfm["rho", 1] *
      x ^ (pfm["rho", 1] - 1),

```

```

    add = TRUE, lty = 4)
curve(qgamma(.05, shape = 1 / pfm["theta", 1],
           scale = pfm["theta", 1]) * exp(pfm["Heifer", 1]) *
      (365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1] * pfm["rho", 1] *
      x ^ (pfm["rho", 1] - 1),
    add = TRUE, lty = 4)

### - Cumulative distribution functions - ###
# multiparous cows
curve(1 - exp(
  -(365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1] *
  x ^ (pfm["rho", 1])),
  from = 0, to = 400, ylim = c(0, 1),
  main = "Multiparous cows",
  ylab = "Cumulative distribution function",
  xlab = "Time to first insemination (days)")
curve(1 - exp(
  -qgamma(.95, shape = 1 / pfm["theta", 1],
           scale = pfm["theta", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
  pfm["lambda", 1] * x ^ (pfm["rho", 1])),
  add = TRUE, lty = 4)
curve(1 - exp(
  -qgamma(.05, shape = 1 / pfm["theta", 1],
           scale = pfm["theta", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
  pfm["lambda", 1] * x ^ (pfm["rho", 1])),
  add = TRUE, lty = 4)

# primiparous cows
curve(1 - exp(
  -exp(pfm["Heifer", 1]) * (365.25 / 12) ^ (-pfm["rho", 1]) *
  pfm["lambda", 1] * x ^ (pfm["rho", 1])),
  from = 0, to = 400, ylim = c(0, 1),
  main = "Primiparous cows",
  ylab = "Cumulative distribution function",
  xlab = "Time to first insemination (days)")
curve(1 - exp(
  -qgamma(.95, shape = 1 / pfm["theta", 1],
           scale = pfm["theta", 1]) * exp(pfm["Heifer", 1]) *
  (365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1] * x ^ (pfm["rho", 1])),
  add = TRUE, lty = 4)
curve(1 - exp(
  -qgamma(.05, shape = 1 / pfm["theta", 1],
           scale=pfm["theta", 1]) * exp(pfm["Heifer", 1]) *
  (365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1] * x ^ (pfm["rho", 1])),
  add = TRUE, lty = 4)

### - Density of the median time - ###
fM <- function(x, heifer = 0) {
  RHO <- pfm["rho", 1]
  LAMBDAad <- (365.25 / 12) ^ (-pfm["rho", 1]) * pfm["lambda", 1]
  THETA <- pfm["theta", 1]

```

```

if (heifer) {
  eBETA <- exp(pfm["Heifer", 1])
} else eBETA <- 1

RHO * (log(2) / (THETA * LAMBDAd * eBETA)) ^ (1 / THETA) *
x^(-1 - RHO / THETA) *
exp(-log(2) / (THETA * x^RHO * LAMBDAd * eBETA)) /
gamma(1 / THETA)
}

par(mfrow=c(1, 1))
curve(fM, 0, 300, xlab = "Median time to first insemination (days)",
      ylab = "Density function of the median")
curve(fM(x, heifer = 1), add = TRUE, lty = 3)
legend("topright", legend = c("Multiparous", "Primiparous"),
      lty = c(1, 3), bty = "n")

```

insemtvc	<i>Time to first insemtvcination in dairy heifer cows with time varying covariates</i>
----------	--

Description

In a dairy farm, the calving interval (the time between two calvings) should be optimally between 12 and 13 months. One of the main factors determining the length of the calving interval is the time from parturition to the time of first insemtvcination.

The objective of this study is to look for cow factors that might predict the time to first insemtvcination, so that actions can be taken based on these predictors. As no insemtvcinations take place in the first 29 days after calving, we subtract 29 days (and not 30 days as the first event would then have first insemtvcination time zero) from the time to first insemtvcination since at risk time starts only then. Cows which are culled without being inseminated are censored at their culling time. Furthermore, cows that are not yet inseminated 300 days after calving are censored at that time.

The time to first insemtvcination is studied in dairy cows as a function of time-varying covariates. Examples of such covariates are the protein and ureum concentration. The protein and ureum concentrations are measured, during the experiment, at a number of points in time. It might be more adequate to model the hazard at a particular time using the concentration at that particular point in time. In order to accommodate for time-varying covariates, the risk time for each cow is split into time intervals with a start and an end time and in each such interval a constant value for the concentration. We therefore have for each cow as many data lines as there are risk intervals.

Usage

```
data(insemtvc)
```

Format

A dataframe containing 10513 observations.

- Cowid:** Cow's identifier.
- Start:** Start time of the interval (in days).
- End:** End time of the interval (in days).
- Status:** Censored (0) or observed (1) event time.
- Herd:** The herd to which the cow belongs.
- Urem:** Milk urem concentration (%) at the begin time of the interval.
- Protein:** Protein concentration (%) at the begin time of the interval.
- Parity:** The number of calvings.
- Heifer:** Multiparous cow (0) or primiparous cow (1).

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.8 of Duchateau an Janssen (2008)

References

Duchateau L., Janssen P. (2008). *The frailty model*. Springer. New York: Springer-Verlag.

mastitis

Correlated infection times in four cow udder quarters

Description

Mastitis, the infection of the udder, is economically the most important disease in the dairy sector of the western world. Mastitis can be caused by many organisms, most of them bacteria, such as *Escherichia coli*, *Streptococcus uberis*, and *Staphylococcus aureus*. Since each udder quarter is separated from the three other quarters, one quarter might be infected with the other quarters free of infection.

In an extensive study, 100 cows are followed up for infections. The objective of this observational study is to estimate the incidence of the different organisms causing mastitis in the dairy cattle population in Flanders. Also the correlation between the infection times of the four udder quarters of a cow is an important parameter to take preventive measures against mastitis. With high correlation, a lot of attention should be given to the uninfected udder quarters of a cow that has an infected quarter. From each quarter, a milk sample is taken monthly and is screened for the presence of different bacteria.

We model the time to infection with any bacteria, with the cow being the cluster and the quarter the experimental unit within the cluster. Observations can be right-censored if no infection occurs before the end of the lactation period, which is roughly 300-350 days but different for every cow, or if the cow is lost to follow-up during the study, for example due to culling. Due to the periodic

follow-up, udder quarters that experience an event are interval-censored with lowerbound the time of the last milk sample with a negative result and upperbound the time of the first milk sample with a positive result. In some examples, the midpoint (average of lowerbound and upperbound of the interval) is used for simplicity; in other examples the interval-censored nature of the data is taken into account.

In the analysis, two types of covariates are considered. Cow level covariates take the same value for every udder quarter of the cow (e.g., number of calvings or parity). Several studies have shown that prevalence as well as incidence of intramammary infections increase with parity. Several hypotheses have been suggested to explain these findings, e.g., teat end condition deteriorates with increasing parity. Because the teat end is a physical barrier that prevents organisms from invading the udder, impaired teat ends make the udder more vulnerable for intramammary infections. For simplicity, parity is dichotomised into primiparous cows (heifer=1) and multiparous cows (heifer=0). Udder quarter level covariates change within the cow (e.g., position of the udder quarter, front or rear). The difference in teat end condition between front and rear quarters has also been put forward to explain the difference in infection status.

Usage

```
data(mastitis)
```

Format

A dataframe containing 400 observations.

Cowid: Cow's identifier.

Lower: Lower bound of time to infection (in days).

Upper: Upper bound of time to infection (in days).

Midpoint: Midpoint between Lower and Upper (in days).

Status: Censored (0) or observed (1) event time.

Heifer: Multiparous cow (0) or primiparous cow (1).

Quarter: The udder quarter (LF=Left-Front, LR=Left-Rear, RF=Right-Front, RR=Right-Rear).

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.4 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```

data(mastitis)
head(mastitis)

mastitis$timeto <- as.numeric((mastitis$Midpoint * 4 / 365.25))

#####
# Example 4.4: The gamma frailty model for the udder quarter infection data #
# Duchateau and Janssen (2008, page 136) #
#####
modParfm <- parfm(Surv(timeto, Status) ~ Heifer,
                  cluster = "Cowid",
                  dist = "weibull",
                  frailty = "gamma",
                  data = mastitis)

modParfm

#####
# Example 4.7 The inverse Gaussian frailty model for the udder quarter #
# infection data #
# Duchateau and Janssen (2008, page 156) #
#####
mastitis <- data.frame(mastitis,
                      timeto = as.numeric((mastitis$Midpoint * 4 / 365.25)))
modParfm <- parfm(Surv(timeto, Status) ~ Heifer,
                  cluster = "Cowid",
                  dist = "weibull",
                  frailty = "ingau",
                  data = mastitis)

modParfm

#####
# Example 4.10 The positive stable frailty model for the udder quarter #
# infection data #
# Duchateau and Janssen (2008, page 169) #
#####
mastitis <- data.frame(mastitis,
                      timeto = as.numeric((mastitis$Midpoint * 4 / 365.25)))
modParfm <- parfm(Surv(timeto, Status) ~ Heifer,
                  cluster = "Cowid",
                  dist = "weibull",
                  frailty = "possta",
                  data = mastitis)

modParfm

```

Description

Fits parametric Frailty Models by maximizing the marginal likelihood.

Usage

```
parfm(formula, cluster = NULL, strata = NULL, data,
      inip = NULL, iniFpar = NULL,
      dist = c("weibull", "inweibull", "frechet", "exponential",
              "gompertz", "loglogistic", "lognormal", "logskewnormal"),
      frailty = c("none", "gamma", "ingau", "possta",
                 "lognormal", "loglogistic"),
      method = "nlminb",
      maxit = 500, Fparscale = 1, showtime = FALSE, correct = 0)
```

Arguments

formula	A formula object, with the response on the left of a ~ operator, and the terms on the right. The response must be a survival object as returned by the Surv() function. The status indicator 'event' in the Surv object must be 0=alive, 1=dead.
cluster	The name of a cluster variable in data.
strata	The name of a strata variable in data.
data	A data.frame in which to interpret the variables named in the formula.
inip	The vector of initial values. First components are for the baseline hazard parameters according to the order given in 'details'; Other components are for the regression parameters according to the order given in 'formula'.
iniFpar	The initial value of the frailty parameter.
dist	The baseline hazard distribution. One of weibull, inweibull, frechet, exponential, gompertz, lognormal, loglogistic, or logskewnormal.
frailty	The Frailty distribution. One of: none, gamma, ingau, possta or lognormal.
method	The optimisation method from the function optimx() .
maxit	Maximum number of iterations (see optimx()).
Fparscale	the scaling value for the frailty parameter in optimx() . Optimisation is performed on Fpar/Fparscale.
showtime	Show the execution time?
correct	A correction factor that does not change the marginal log-likelihood except for an additive constant given by #clusters * correct * log(10). It may be useful in order to get finite log-likelihood values in case of many events per cluster with Positive Stable frailties. Note that the value of the log-likelihood in the output is the re-adjusted value.

Details**Baseline hazards**

The Weibull hazard (dist="weibull") is

$$h(t; \rho, \lambda) = \rho \lambda t^{\rho-1}$$

with $\rho, \lambda > 0$.

The inverse Weibull (or Fréchet) hazard (dist="inweibull" or dist="frechet") is

$$h(t; \rho, \lambda) = \frac{\rho \lambda t^{-\rho-1}}{\exp(\lambda t^{-\rho}) - 1}$$

with $\rho, \lambda > 0$.

The exponential hazard (dist="exponential") is

$$h(t; \lambda) = \lambda$$

with $\lambda > 0$.

The Gompertz hazard (dist="gompertz") is

$$h(t; \gamma, \lambda) = \lambda \exp(\gamma t)$$

with $\gamma, \lambda > 0$.

The lognormal hazard (dist="lognormal") is

$$h(t; \mu, \sigma) = \frac{\phi\left(\frac{\log(t)-\mu}{\sigma}\right)}{\sigma t \left[1 - \Phi\left(\frac{\log(t)-\mu}{\sigma}\right)\right]}$$

with $\mu \in R, \sigma > 0$, and where ϕ and Φ are the density and distribution functions of a standard Normal random variable.

The loglogistic hazard (dist="loglogistic") is

$$h(t; \alpha, \kappa) = \frac{\exp(\alpha) \kappa t^{\kappa-1}}{1 + \exp(\alpha) t^{\kappa}}$$

with $\alpha \in R, \kappa > 0$.

The log-skew-normal hazard (dist="logskewnormal") is obtained as the ratio between the density and the cumulative distribution function of a log-skew normal random variable (Azzalini, 1985), which has density

$$f(t; \xi, \omega, \alpha) = \frac{2}{\omega t} \phi\left(\frac{\log(t) - \xi}{\omega}\right) \Phi\left(\alpha \frac{\log(t) - \xi}{\omega}\right)$$

with $\xi \in R, \omega > 0, \alpha \in R$, and where ϕ and Φ are the density and distribution functions of a standard Normal random variable. Of note, if $\alpha = 0$ then the log-skew-normal boils down to the log-normal distribution, with $\mu = \xi$ and $\sigma = \omega$.

Frailty distributions

The gamma frailty distribution (frailty="gamma") is

$$f(u) = \frac{\theta^{-\frac{1}{\theta}} u^{\frac{1}{\theta}-1} \exp(-u/\theta)}{\Gamma(1/\theta)}$$

with $\theta > 0$ and where $\Gamma(\cdot)$ is the gamma function.

The inverse Gaussian frailty distribution (frailty="ingau") is

$$f(u) = \frac{1}{\sqrt{2\pi\theta}} u^{-\frac{3}{2}} \exp\left(-\frac{(u-1)^2}{2\theta u}\right)$$

with $\theta > 0$.

The positive stable frailty distribution (frailty="poosta") is

$$f(u) = f(u) = -\frac{1}{\pi u} \sum_{k=1}^{\infty} \frac{\Gamma(k(1-\nu)+1)}{k!} (-u^{\nu-1})^k \sin((1-\nu)k\pi)$$

with $0 < \nu < 1$.

The lognormal frailty distribution (frailty="lognormal") is

$$f(u) = \frac{1}{\sqrt{2\pi\theta}} u^{-1} \exp\left(-\frac{\log(u)^2}{2\theta}\right)$$

with $\theta > 0$. As the Laplace transform of the lognormal frailties does not exist in closed form, the saddlepoint approximation is used (Goutis and Casella, 1999).

Value

An object of class parfm.

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Munda M, Rotolo F, Legrand C (2012). parfm: Parametric Frailty Models in R. Journal of Statistical Software, 51(11), 1-20. DOI <doi: [10.18637/jss.v051.i11](https://doi.org/10.18637/jss.v051.i11)>

Duchateau L, Janssen P (2008). The frailty model. Springer.

Wienke A (2010). Frailty Models in Survival Analysis. Chapman & Hall/CRC biostatistics series. Taylor and Francis.

Goutis C, Casella G (1999). Explaining the Saddlepoint Approximation. The American Statistician 53(3), 216-224. DOI <doi: [10.1080/00031305.1999.10474463](https://doi.org/10.1080/00031305.1999.10474463)>

Azzalini A (1985). A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12(2):171-178.

See Also

[select.parfm](#), [ci.parfm](#), [predict.parfm](#)

Examples

```

#-----Kidney dataset-----
data(kidney)
# type 'help(kidney)' for a description of the data set
kidney$sex <- kidney$sex - 1

parfm(Surv(time,status) ~ sex + age, cluster = "id",
      data = kidney, dist = "exponential", frailty = "gamma")

parfm(Surv(time,status) ~ sex + age, cluster = "id",
      data = kidney, dist = "exponential", frailty = "lognormal")

parfm(Surv(time,status) ~ sex + age, cluster = "id",
      data = kidney, dist = "weibull", frailty = "ingau")

parfm(Surv(time,status) ~ sex + age, cluster = "id",
      data = kidney, dist="gompertz", frailty="possta", method="CG")

parfm(Surv(time,status) ~ sex + age, cluster = "id",
      data = kidney, dist="logskewnormal", frailty="possta", method = 'BFGS')
#-----

#-----Asthma dataset-----
data(asthma)
head(asthma)
# type 'help(asthma)' for a description of the data set

asthma$time <- asthma$End - asthma$Begin
parfm(Surv(time, Status) ~ Drug, cluster = "Patid", data = asthma,
      dist = "weibull", frailty = "gamma")

parfm(Surv(time, Status) ~ Drug, cluster = "Patid", data = asthma,
      dist = "weibull", frailty = "lognormal")

parfm(Surv(Begin, End, Status) ~ Drug, cluster = "Patid",
      data = asthma[asthma$Fevent == 0, ],
      dist = "weibull", frailty = "lognormal", method = "Nelder-Mead")
#-----

```

predict.parfm

Predictions of frailty values for Parametric Frailty Models

Description

The function `predict.parfm()` computes predictions of frailty values for objects of class `parfm`.

Usage

```
## S3 method for class 'parfm'
predict(object, ...)
```

Arguments

```
object      A parametric frailty model, object of class parfm.
...         see predict\(\)
```

Value

An object of class `predict.parfm`.

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Glidden D, Vittinghoff E (2004). Modelling Clustered Survival Data From Multicentre Clinical Trials. *Statistics in medicine*, 23(3), 369–388.

Munda M, Rotolo F, Legrand C (2012). `parfm`: Parametric Frailty Models in R. *Journal of Statistical Software*, 51(11), 1-20. DOI <doi: [10.18637/jss.v051.i11](https://doi.org/10.18637/jss.v051.i11)>

See Also

[parfm](#)

Examples

```
data(kidney)
kidney$sex <- kidney$sex - 1

model <- parfm(Surv(time,status) ~ sex + age,
              cluster = "id", data = kidney,
              dist = "exponential", frailty = "gamma")
u <- predict(model)
u

# Predictions from semi-parametric Gamma frailty model
# via coxph() function
model.coxph <- coxph(Surv(time,status) ~ sex + age +
                    frailty(id, frailty = "gamma", eps = 1e-11),
                    outer.max = 15, data = kidney)
u.coxph <- exp(model.coxph$frail)

# Plot of predictions from both models
par(mfrow = c(1,2))
```

```

ylim <- c(0, max(c(u, u.coxph)))
plot(u, sort = "i",
      main = paste("Parametric",
                  "Gamma frailty model",
                  "with Exponential baseline",
                  sep = "\n"),
      ylim = ylim)

names(u.coxph) <- kidney[seq(2,76, 2), "id"]
class(u.coxph) <- "predict.parfm"
attr(u.coxph, "clustname") <- "id"
plot(u.coxph, sort = "i",
      main = paste("Semi-parametric",
                  "Gamma frailty model", sep = "\n"),
      ylim = ylim)

```

reconstitution

Reconstitution of blood–milk barrier after reconstitution

Description

When an udder quarter of a cow is infected (reconstitution), the blood-milk barrier is partially destroyed and particular ions can flow freely from blood to milk and vice versa, leading to higher concentrations of, for instance, the sodium concentration Na.

The objective of this study is to demonstrate that the local application of a drug based on corticosteroids decreases the time to reconstitution of the blood-milk barrier in dairy cows. We therefore consider as outcome the time until the Na concentration goes below a certain threshold (a concentration below the threshold value is considered to be normal again). Each udder quarter is separated from the three other quarters so that a quarter can be used as experimental unit to which a treatment is assigned. The Na concentration in each of the experimental units is followed up. The rear udder quarters of 100 cows are experimentally infected with *Escherichia coli*. After nine hours, one of the two infected udder quarters is treated locally with the active compound whereas the other is treated with placebo. Cows are followed up for 6.5 days, and are censored at that point in time if the Na concentration is still above the threshold level.

We further include parity in the study as covariate. The parity of a cow is the number of calvings (and therefore the number of lactation periods) that the cow has already experienced. Parity is often converted into a binary covariate, grouping all the cows with more than one calving in the group of multiparous cows (`heifer=0`) compared to the group of primiparous cows or heifers, cows with only one calving (`heifer=1`).

Usage

```
data(reconstitution)
```

Format

A dataframe containing 200 observations.

Cowid: Cow's identifier.

Time: The time to reconstitution (in days).

Status: Censored (0) or observed (1) event time.

Drug: Drug or placebo application.

Heifer: Multiparous cow (0) or primiparous cow (1).

Note

These data simulated, with exactly the same structure as the real data used in the book, that could not be made publicly available.

Source

Example 1.3 of Duchateau and Janssen (2008)

References

Duchateau L, Janssen P (2008). *The frailty model*. Springer. New York: Springer-Verlag.

Examples

```
data(reconstitution)
head(reconstitution)
```

```
#####
# Example 3.1: The [...] frailty model for the time to blood-milk barrier      #
# reconstitution: the effect of treatment                                    #
# Duchateau and Janssen (2008, page 79)                                     #
#####
pfmDrug <- parfm(Surv(Time, Status) ~ Drug,
                 cluster = "Cowid", dist = "exponential",
                 frailty = "gamma", data = reconstitution)
ci.parm(pfmDrug)
```

```
#####
# Example 3.2: The [...] frailty model for the time to blood-milk barrier      #
# reconstitution: the heifer effect                                          #
# Duchateau and Janssen (2008, page 82)                                       #
#####
pfmHeifer <- parfm(Surv(Time, Status) ~ Heifer,
                  cluster = "Cowid", dist = "exponential",
                  frailty = "gamma", data = reconstitution)
ci.parm(pfmHeifer)
```

select.parfm *AIC and BIC values of several Parametric Frailty Models*

Description

The function `select.parfm()` computes the AIC and BIC values of parametric frailty models with different baseline hazards and different frailty distributions.

Usage

```
select.parfm(formula, cluster=NULL, strata=NULL, data, inip=NULL, iniFpar=NULL,
             dist=c("exponential", "weibull", "inweibull", "frechet", "gompertz",
                   "loglogistic", "lognormal", "logskewnormal"),
             frailty=c("none", "gamma", "ingau", "possta", "lognormal"),
             method="BFGS", maxit=500, Fparscale=1, correct=0)
```

Arguments

formula	A formula object, with the response on the left of a ~ operator, and the terms on the right. The response must be a survival object as returned by the Surv function.
cluster	The name of a cluster variable in data.
strata	The name of a strata variable in data.
data	A data.frame in which to interpret the variables named in the formula.
inip	The vector of initial values. First components are for the baseline hazard parameters according to the order given in 'details'; Other components are for the regression parameters according to the order given in 'formula'.
iniFpar	The initial value of the frailty parameter.
dist	The vector of baseline hazards' names. It can include any of <code>weibull</code> , <code>inweibull</code> (alias <code>frechet</code>), <code>exponential</code> , <code>gompertz</code> , <code>loglogistic</code> or <code>lognormal</code> .
frailty	The vector of frailty distributions' names. It can include any of: <code>none</code> , <code>gamma</code> , <code>ingau</code> , <code>possta</code> or <code>lognormal</code> .
method	The optimisation method from the function optim() .
maxit	Maximum number of iterations (see optim()).
Fparscale	the scaling value for the frailty parameter in optim() . Optimisation is performed on <code>Fpar/Fparscale</code> .
correct	A correction factor that does not change the marginal log-likelihood except for an additive constant given by <code>#clusters * correct * log(10)</code> . It may be useful in order to get finite log-likelihood values in case of many events per cluster with Positive Stable frailties. Note that the value of the log-likelihood in the output is the re-adjusted value.

Value

An object of class `select.parfm`.

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Munda M, Rotolo F, Legrand C (2012). parfm: Parametric Frailty Models in R. Journal of Statistical Software, 51(11), 1-20. DOI <doi: [10.18637/jss.v051.i11](https://doi.org/10.18637/jss.v051.i11)>

See Also

[parfm](#), [ci.parfm](#), [predict.parfm](#)

Examples

```
data(kidney)
kidney$sex <- kidney$sex - 1

models <- select.parfm(Surv(time,status) ~ sex + age,
  dist = c("exponential",
           "weibull",
           "inweibull",
           "loglogistic",
           "lognormal",
           "logskewnormal"),
  frailty = c("gamma",
             "ingau",
             "possta",
             "lognormal"),
  cluster = "id", data = kidney)

models
plot(models)
```

tau

Kendall's Tau for Parametric Frailty Models

Description

Computes Kendall's Tau for Parametric Frailty Models

Usage

```
tau(x)
```

Arguments

x A parametric frailty model, object of class [parfm](#).

Author(s)

Federico Rotolo [aut, cre], Marco Munda [aut], Andrea Callegaro [ctb]

References

Kendall, MG. (1938) A new measure of rank correlation. *Biometrika* 30(1/2), 81-93. DOI: <doi: [10.2307/2332226](https://doi.org/10.2307/2332226)>

See Also

[parfm](#)

Examples

```
data(kidney)
# type 'help(kidney)' for a description of the data set
kidney$sex <- kidney$sex - 1

mod <- parfm(Surv(time,status) ~ sex + age, cluster = "id",
             data = kidney, dist = "exponential", frailty = "gamma")
tau(mod)
```

Index

- * **East Coast Fever**
 - ecf, 9
- * **Kendall**
 - tau, 26
- * **RX**
 - diagnosis, 7
- * **Theileriosis**
 - ecf, 9
- * **US**
 - diagnosis, 7
- * **asthma**
 - asthma, 3
- * **cows**
 - culling, 6
 - ecf, 9
 - insem, 11
 - insemtvc, 14
 - mastitis, 15
 - reconstitution, 23
- * **culling**
 - culling, 6
- * **datasets**
 - asthma, 3
 - culling, 6
 - diagnosis, 7
 - ecf, 9
 - insem, 11
 - insemtvc, 14
 - mastitis, 15
 - reconstitution, 23
- * **diagnosis**
 - diagnosis, 7
- * **dogs**
 - diagnosis, 7
- * **ecf**
 - ecf, 9
- * **fracture**
 - diagnosis, 7
- * **frailty prediction**
 - predict.parfm, 21
- * **frailty**
 - asthma, 3
 - culling, 6
 - diagnosis, 7
 - ecf, 9
 - insem, 11
 - insemtvc, 14
 - mastitis, 15
 - parfm, 17
 - predict.parfm, 21
 - reconstitution, 23
 - tau, 26
- * **infection**
 - ecf, 9
 - mastitis, 15
 - reconstitution, 23
- * **insemination**
 - insem, 11
- * **insemtvcination**
 - insemtvc, 14
- * **insemtvc**
 - insemtvc, 14
- * **insem**
 - insem, 11
- * **marginal likelihood**
 - parfm, 17
 - tau, 26
- * **mastitis**
 - mastitis, 15
- * **models**
 - anova.parfm, 2
- * **parametric**
 - parfm, 17
 - tau, 26
- * **prediction**
 - predict.parfm, 21
- * **radiography**
 - diagnosis, 7

- * **reconstitution**
 - reconstitution, 23
 - * **regression**
 - anova.parfm, 2
 - * **shared**
 - parfm, 17
 - tau, 26
 - * **survival**
 - anova.parfm, 2
 - parfm, 17
 - tau, 26
 - * **tau**
 - tau, 26
 - * **udder**
 - mastitis, 15
 - reconstitution, 23
 - * **ultrasound**
 - diagnosis, 7
- anova, 3
- anova.parfm, 2
- anova.parfmlist (anova.parfm), 2
- asthma, 3
- ci.parfm, 5, 20, 26
- coef.parfm (parfm), 17
- culling, 6
- data.frame, 18, 25
- diagnosis, 7
- ecf, 9
- formula, 18, 25
- insem, 11
- insemtvc, 14
- logLik.parfm (parfm), 17
- mastitis, 15
- optim, 25
- optimx, 18
- parfm, 3, 5, 17, 21, 22, 26, 27
- plot.parfm (parfm), 17
- plot.predict.parfm (predict.parfm), 21
- plot.select.parfm (select.parfm), 25
- predict, 22
- predict.parfm, 5, 20, 21, 26
- print.parfm (parfm), 17
- print.predict.parfm (predict.parfm), 21
- print.select.parfm (select.parfm), 25
- reconstitution, 23
- select.parfm, 5, 20, 25
- skew normal, 19
- Surv, 18, 25
- tau, 26