

Package: afthd (via r-universe)

August 30, 2024

Title Accelerated Failure Time for High Dimensional Data with MCMC

Type Package

Version 1.1.0

Date 2021-10-01

Description Functions for Posterior estimates of Accelerated Failure Time(AFT) model with MCMC and Maximum likelihood estimates of AFT model without MCMC for univariate and multivariate analysis in high dimensional gene expression data are available in this 'afthd' package. AFT model with Bayesian framework for multivariate in high dimensional data has been proposed by Prabhash et al.(2016) <[doi:10.21307/stattrans-2016-046](https://doi.org/10.21307/stattrans-2016-046)>.

LazyDataCompression xz

ByteCompile Yes

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.5.0)

Imports glmnet,photobiology,R2jags,rstpm2,survival

Maintainer Atanu Bhattacharjee <atanustat@gmail.com>

RoxygenNote 7.1.2

NeedsCompilation no

Date/Publication 2021-10-01 10:00:05 UTC

Author Atanu Bhattacharjee [aut, cre, ctb], Gajendra Kumar Vishwakarma [aut, ctb], Pragyaa Kumari [aut, ctb]

Repository <https://atanubhattacharjee.r-universe.dev>

RemoteUrl <https://github.com/cran/afthd>

RemoteRef HEAD

RemoteSha 5ec629860e84581615cb17f130a7081bb1eadc1d

Contents

aftbybmv	2
hdata	4
lgnbymv	4
lgnbyuni	6
lgstbymv	7
lgstbyuni	8
pvaft	10
rglaft	11
rglwbyism	12
rglwbyisu	14
wbyAgmv	15
wbyscrkm	16
wbyscrku	18
wbysmv	19
wbysuni	20
Index	22

aftbybmv	<i>Bayesian multivariate analysis of parametric AFT model with minimum deviance (DIC) among weibull, log normal and log logistic distribution.</i>
----------	--

Description

Provides better estimates (which has minimum deviance(DIC)) for survival data among weibull, log normal and log logistic distribution of parametric AFT model using MCMC for multivariable (maximum 5 at a time) in high dimensional data.

Usage

```
aftbybmv(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of MCMC chain.
ni	number of MCMC iteration to update the outcome.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x_i' \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

i.e.,

$$T_i \sim AFT(F_\epsilon, \beta, \tau | x_i)$$

Where F_ϵ is known cdf which is defined on real line. Here, when baseline distribution is extreme value then T follows weibull distribution. To make interpretation of regression coefficients simpler, using extreme value distribution with median 0. So using weibull distribution that leads to AFT model when

$$T \sim Weib(\sqrt{\tau}, \log(2) \exp(-x' \beta \sqrt{\tau}))$$

When baseline distribution is normal then T follows log normal distribution.

$$T \sim LN(x' \beta, 1/\tau)$$

When baseline distribution is logistic then T follows log logistic distribution.

$$T \sim Log - Logis(x' \beta, \sqrt{\tau})$$

Value

Data frame is containing posterior estimates mean, sd, credible intervals, n.eff and Rhat for beta's, sigma, tau and deviance of the model for the selected covariates. beta's of regression coefficient of the model. beta[1] is for intercept and others are for covariates (which is/are chosen order as columns in data). 'sigma' is the scale parameter of the distribution. DIC is the estimate of expected predictive error (so lower deviance denotes better estimation).

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysmv, lgnbmv, lgstbmv

Examples

```
##
data(hdata)
aftbybmv(10, 12, STime="os", Event="death", 2, 100, hdata)
##
```

hdata	<i>Head and neck cancer data</i>
-------	----------------------------------

Description

High dimensional head and neck cancer gene expression data

Usage

```
data(hdata)
```

Format

A dataframe with 565 rows and 104 variables

ID ID of subjects

leftcensoring Initial censoring time

death Survival event

death2 death due to other causes

os Duration of overall survival

PFS Duration of progression free survival

Prog Progression event

GJB1, ..., HMGCS2 High dimensional covariates

Examples

```
data(hdata)
```

lgnbymv	<i>Bayesian multivariate analysis of AFT model with log normal distribution.</i>
---------	--

Description

Provides posterior estimates of AFT model with log normal distribution using Bayesian for multivariate (maximum 5 at a time) in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
lgnbymv(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of MCMC chain.
ni	number of MCMC iteration to update the outcome.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x_i' \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

Where F_ϵ is known cdf which is defined on real line. When baseline distribution is normal then T follows log normal distribution.

$$T \sim LN(x' \beta, 1/\tau)$$

Value

Data frame is containing mean, sd, n.eff, Rhat and credible intervals for beta's, sigma, tau and deviance of the model for the chosen covariates. beta[1] is for intercept and others are for covariates (which is/are chosen as columns in data). sigma is the scale parameter of the distribution.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

lgnbyuni, wbysmv, lgstbymv

Examples

```
##
data(hdata)
lgnbymv(10, 12, STime="os", Event="death", 2, 100, hdata)
##
```

lgnbyuni	<i>Bayesian univariate analysis of AFT model with log normal distribution.</i>
----------	--

Description

Provides posterior estimates of AFT model with log normal distribution using Bayesian for univariate in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
lgnbyuni(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of MCMC chain.
ni	number of MCMC iteration to update the outcome.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x_i' \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

Where F_ϵ is known cdf which is defined on real line. When baseline distribution is normal then T follows log normal distribution.

$$T \sim LN(x' \beta, 1/\tau)$$

Value

Data frame is containing posterior estimates (Coef, SD, Credible Interval, Rhat, n.eff) of regression coefficient of selected covariates and deviance. Result shows together for all covariates chosen from column m to n.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

lgnbmv, wbysuni, lgstbyuni

Examples

```
##
data(hdata)
lgnbyuni(10,12,STime="os",Event="death",2,10,hdata)
##
```

lgstbmv	<i>Multivariate estimates of AFT model with log logistic distribution using MCMC.</i>
---------	---

Description

Provides estimate of AFT model with log logistic distribution using MCMC for multivariable (maximum 5 covariates of column at a time) in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
lgstbmv(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of MCMC chain.
ni	number of MCMC iteration to update the outcome.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x_i' \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

Where F_ϵ is known cdf which is defined on real line. When baseline distribution is logistic then T follows log logistic distribution.

$$T \sim \text{Log} - \text{Logis}(x' \beta, \sqrt{\tau})$$

Value

Data frame is containing mean, sd, n.eff, Rhat and credible intervals (2.5%, 25%, 50%, 75% and 97.5%) for beta's, sigma, tau and deviance of the model for the selected covariates. beta[1] is for intercept and others are for covariates (which is/are chosen as columns in data). sigma is the scale parameter of the distribution.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi: 10.21307/stattrans-2016-046>

See Also

wbysmv, lgnbymv, lgstbyuni

Examples

```
##
data(hdata)
lgstbymv(10, 12, STime="os", Event="death", 5, 100, hdata)
##
```

 lgstbyuni

Univariate estimates of AFT model with log logistic distribution using MCMC.

Description

Provides estimate of AFT model with log logistic distribution using MCMC for univariate in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
lgstbyuni(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event in data
nc	number of chain used in model.
ni	number of iteration used in model.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x'_i \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

Where F_ϵ is known cdf which is defined on real line. When baseline distribution is logistic then T follows log logistic distribution.

$$T \sim \text{Log} - \text{Logis}(x' \beta, \sqrt{\tau})$$

Value

Data frame is containing posterior estimates (Coef, SD, Credible Interval, Rhat, n.eff) of regression coefficient of selected covariates and deviance. Result shows together for all covariates chosen from column m to n.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysmv, lgnbymv, lgstbymvs

Examples

```
##
data(hdata)
lgstbyuni(12,14,STime="os",Event="death",3,100,hdata)
##
```

pvaft *Estimates of univariate covariates using Accelerated Failure time (AFT) model without MCMC.*

Description

Provides list of covariates and their estimates of parametric AFT model with smooth time functions, whose p value is less than chosen value (by default p=1 that is all chosen covariates come in result). Using AFT model for univariate in high dimensional data without MCMC.

Usage

```
pvaft(m, n, STime, Event, p = 1, data)
```

Arguments

m	Starting column number of covariates of study in high dimensional entered data.
n	Ending column number of covariates of study in high dimensional entered data.
STime	name of survival time in data.
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
p	p-value, to make restriction for selection of covariates, default value is 1.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

Survival time T for covariate x, is modelled as AFT model using

$$S(T|x) = S_0(T \exp(-\eta(x; \beta)))$$

and baseline survival function is modelled as

$$S_0(T) = \exp(-\exp(\eta_0(\log(T); \beta_0)))$$

Where η and η_0 are linear predictor.

Value

Matrix that contains survival information of selected covariates(selected from chosen columns whose p value is $\leq p$) on AFT model. Result shows together for all covariates chosen from column m to n.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

See Also

wbysuni,wbysmv,rglaft

Examples

```
##
data(hdata)
pvaft(9,30,STime="os",Event="death",0.1,hdata)
##
```

rglaft	<i>Estimates of selected univariate covariates(using regularization) in AFT model without MCMC.</i>
--------	---

Description

Provides Estimates of selected variable in parametric AFT model with smooth time functions for univariate in high dimensional gene expression data without MCMC. Incorporated variable selection has been done with regularization technique. It also deals covariates with missing values.

@details Survival time T for covariate x , is modelled as AFT model using

$$S(T|x) = S_0(T \exp(-\eta(x; \beta)))$$

and baseline survival function is modelled as

$$S_0(T) = \exp(-\exp(\eta_0(\log(T); \beta_0)))$$

Where η and η_0 are linear predictor.

Usage

```
rglaft(m, n, STime, Event, alpha, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
alpha	It is chosen value between 0 and 1 to know the regularization method. alpha=1 for Lasso, alpha=0 for Ridge and alpha between 0 and 1 for elastic net regularization.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Value

Matrix that contains survival information of selected covariates(selected from chosen columns using regularization) on AFT model. Uppermost covariates are more significant than lowerone, as covariates are ordered as their increasing order of p value.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

See Also

pvaft, rglwbysu, rglwbysm

Examples

```
##
data(hdata)
set.seed(1000)
rglaft(9,50,STime="os",Event="death",1,hdata)
##
```

rglwbysm

Bayesian multivariate analysis of AFT model for selected covariates using regularization method.

Description

Provides posterior Estimates of selected variable in AFT model for multivariable(maximum 5 at a time) in high dimensional gene expression data with MCMC.Incorporated variable selection has been done with regularization technique. It also deals covariates with missing values.

Usage

```
rglwbysm(m, n, STime, Event, nc, ni, alpha, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event status in data. 0 is for censored and 1 for occurrence of event.
nc	number of markov chain.
ni	number of iteration for MCMC.

alpha	It is chosen value between 0 and 1 to know the regularization method. alpha=1 for Lasso, alpha=0 for Ridge and alpha between 0 and 1 for elastic net regularization.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

Here weibull distribution has been used for AFT model with MCMC. This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate.

Value

Data frame is containing posterior estimates mean, sd, credible intervals, n.eff and Rhat for beta's, sigma, alpha, tau and deviance (DIC information) of the model for the selected covariates using regularization technique. beta's of regression coefficient of the model. beta[1] is for intercept and others are for covariates (which is/are chosen order as columns in data). alpha is shape parameter of the distribution. 'sigma' is the scale parameter of the distribution.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016)<doi:10.21307/stattrans-2016-046>

See Also

wbysuni, wbysmv, rglwbysu, aftbybmv

Examples

```
##
data(hdata)
set.seed(1000)
rglwbysm(9,45,STime="os",Event="death",2,10,1,hdata)
##
```

rglwbysu	<i>Bayesian univariate analysis of AFT model for selected covariates using regularization method.</i>
----------	---

Description

Provides posterior Estimates of selected variable(by regularization technique) in AFT model for univariate in high dimensional gene expression data with MCMC. Incorporated variable selection has been done with regularization technique. It also deals covariates with missing values.

Usage

```
rglwbysu(m, n, STime, Event, nc, ni, alpha, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of chain used in model.
ni	number of iteration used in model.
alpha	It is chosen value between 0 and 1 to know the regularization method. alpha=1 for Lasso, alpha=0 for Ridge and alpha between 0 and 1 for elastic net regularization.
data	High dimensional gene expression data that contains event status, survival time and set of covariates.

Details

Here weibull distribution has been used for AFT model with MCMC. This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate.

Value

posterior estimates (Coef, SD, Credible Interval) of regression coefficient for all selected covariate (using regularization) in model and deviance.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysuni,wbysmv, rglwbysm

Examples

```
##
data(hdata)
set.seed(1000)
rglwbysu(9,45,STime="os",Event="death",2,10,1,hdata)
##
```

wbyAgmv	<i>Multivariate estimates of AFT model with weibull distribution using MCMC that supports augmented data.</i>
---------	---

Description

Provides estimate of AFT model including Survival time for augmented data with weibull distribution using MCMC for multivariable (maximum 5 covariates of column at a time) in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
wbyAgmv(m, n, p, q, t, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study in data.
n	Ending column number of covariates of study in data.
p	starting row number for augmented data in entered data.
q	last row number for augmented data in entered data
t	time (same unit as STime in data) after which, estimated STime to be printed (for individuals p to q).
STime	name of survival time in data
Event	name of event in data. 0 is for censored and 1 for occurrence of event.
nc	number of markov chain.
ni	number of iteration for MCMC.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

Here weibull distribution has been used for AFT model with MCMC. This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate.

Value

Posterior estimates of beta's, sigma, tau and deviance are their estimates mean, sd, credible intervals, number of efficient sample (n.eff) and Rhat. beta's denotes posterior estimates of regression coefficient of the model. beta[1] is for intercept and others are for covariates (which is/are chosen as columns in data). 'sigma' is the scale parameter of the distribution. 'STime' in output section provides estimated value of STime="os" in data only for individual row number p to q. 'Overall_S' in output, provides an overall estimate of STime="os" in data for all individuals nrow(data). @import R2jags

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysmv

Examples

```
##
data(hdata)
wbyAgmv(9,13,p=560,q=565,t=200,STime="os",Event="death",2,10,hdata)
#
##
```

wbyscrkm

Bayesian multivariate estimates for competing risk gene expression data using AFT model.

Description

Provides multivariate(maximum 5 covariates of column at a time) posterior estimates of AFT model using MCMC for competing risk high dimensional gene expression data. It also deal with missing values.

Usage

```
wbyscrkm(m, n, STime, Event, nc, ni, data)
```


Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event status in data. 0 is for censored and 1 for occurrence of event of interest and 2 for occurrence of event due to other causes.
nc	number of markov chain.
ni	number of iteration for MCMC.
data	High dimensional gene expression data that contains event status with competing risk, survival time and set of covariates.

Details

Here AFT model has been used with weibull distribution. This function deals covariates (in data) with missing values. Missing value in any column(covariate) is replaced by mean of that particular covariate.

Value

Data frame is containing posterior estimates mean, sd, credible intervals, n.eff and Rhat for beta's, sigma, alpha, tau and deviance (DIC information) of the model for the chosen covariates as columns between m and n. beta's of regression coefficient of the model. beta[1] is for intercept and others are for covariates (which is/are chosen order as columns in data). alpha is shape parameter of the distribution. 'sigma' is the scale parameter of the distribution.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysmv, wbyscrku

Examples

```
##
data(hdata)
wbyscrkm(9, 11, STime="os", Event="death2", 2, 10, hdata)
##
```

wbyscrku	<i>Bayesian univariate estimates for competing risk gene expression data using AFT model.</i>
----------	---

Description

Provides univariate estimate of AFT model using MCMC for competing risk high dimensional gene expression data. It also deal with missing values.

Usage

```
wbyscrku(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event status in data. 0 is for censored and 1 for occurrence of event of interest and 2 for occurrence of event due to other causes.
nc	number of markov chain.
ni	number of iteration for MCMC.
data	High dimensional gene expression data that contains event status with competing risk, survival time and and set of covariates.

Details

Here AFT model has been used with weibull distribution. This function deals covariates (in data) with missing values. Missing value in any column(covariate) is replaced by mean of that particular covariate.

Value

posterior estimates (Coef, SD, Credible Interval) of regression coefficient of covariate(which is/are chosen as columns in data) in model and deviance are there. Outcome shows together for all covariates chosen from column m to n.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

wbysuni, wbyscrkm

Examples

```
##
data(hdata)
#1<=p<=q<=nrow(data)
wbyscrku(9,13,STime="os",Event="death2",2,100,hdata)
##
```

wbysmv

Posterior multivariate estimates of AFT model with weibull distribution using MCMC.

Description

Provides estimate of AFT model with weibull distribution using MCMC for multivariable (maximum 5 covariates of column at a time) in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
wbysmv(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data.
Event	name of event status in data. 0 is for censored and 1 for occurrence of event.
nc	number of markov chain.
ni	number of iteration for MCMC.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x_i' \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(\cdot) \text{ which is iid}$$

Where F_e is known cdf which is defined on real line. Here, when baseline distribution is extreme value then T follows weibull distribution. To make interpretation of regression coefficients simpler, using extreme value distribution with median 0. So using weibull distribution that leads to AFT model when

$$T \sim Weib(\sqrt{\tau}, \log(2) \exp(-x' \beta \sqrt{\tau}))$$

Value

Data frame is containing mean, sd, n.eff, Rhat and credible intervals for beta's, sigma, alpha, tau and deviance of the model for the chosen covariates. beta[1] is for intercept and others are for covariates (which is/are chosen as columns in data). sigma is the scale parameter of the distribution. alpha is shape parameter of the distribution.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

pvaft, wbysuni, rglwbysm, wbyscrkm, wbyAgmv

Examples

```
##
data(hdata)
wbysmv(9, 13, STime="os", Event="death", 2, 10, hdata)
##
```

wbysuni

Posterior univariate estimates of AFT model with weibull distribution using MCMC.

Description

Provides posterior estimates of AFT model with weibull distribution using MCMC for univariate in high dimensional gene expression data. It also deals covariates with missing values.

Usage

```
wbysuni(m, n, STime, Event, nc, ni, data)
```

Arguments

m	Starting column number of covariates of study from high dimensional entered data.
n	Ending column number of covariates of study from high dimensional entered data.
STime	name of survival time in data
Event	name of event status in data. 0 is for censored and 1 for occurrence of event.
nc	number of markov chain.
ni	number of iteration for MCMC.
data	High dimensional gene expression data that contains event status, survival time and and set of covariates.

Details

This function deals covariates (in data) with missing values. Missing value in any column (covariate) is replaced by mean of that particular covariate. AFT model is log-linear regression model for survival time T_1, T_2, \dots, T_n . i.e.,

$$\log(T_i) = x'_i \beta + \sigma \epsilon_i; \epsilon_i \sim F_\epsilon(.) \text{ which is iid}$$

Where F_ϵ is known cdf which is defined on real line. Here, when baseline distribution is extreme value then T follows weibull distribution. To make interpretation of regression coefficients simpler, using extreme value distribution with median 0. So using weibull distribution that leads to AFT model when

$$T \sim Weib(\sqrt{\tau}, \log(2) \exp(-x' \beta \sqrt{\tau}))$$

Value

Data frame is containing posterior estimates (Coef, SD, Credible Interval, Rhat, n.eff) of regression coefficient for covariates and deviance. Result shows together for all covariates chosen from column m to n.

Author(s)

Atanu Bhattacharjee, Gajendra Kumar Vishwakarma and Pragya Kumari

References

Prabhash et al(2016) <doi:10.21307/stattrans-2016-046>

See Also

pvaft, wbysmv, rglwbysu, wbyscrku

Examples

```
##
data(hdata)
wbysuni(9,13,STime="os",Event="death",1,10,hdata)
##
```

Index

* datasets

hdata, 4

aftbybmv, 2

hdata, 4

lgnbymv, 4

lgnbyuni, 6

lgstbymv, 7

lgstbyuni, 8

pvaft, 10

rglaft, 11

rglwbyism, 12

rglwbyisu, 14

wbyAgmv, 15

wbyscrkm, 16

wbyscrku, 18

wbysmv, 19

wbysuni, 20