Package 'tailloss'

October 14, 2022

Title Estimate the Probability in the Upper Tail of the Aggregate Loss Distribution

Description Set of tools to estimate the probability in the upper tail of the aggregate loss distribution using different methods: Panjer recursion, Monte Carlo simulations, Markov bound, Cantelli bound, Moment bound, and Chernoff bound.

Version 1.0

Depends R (>= 3.0.2), MASS, graphics, stats

License GPL-2 | GPL-3

LazyData true

URL http://github.com/igollini/tailloss

NeedsCompilation no

Author Isabella Gollini [aut, cre],
Jonathan Rougier [ctb]

Maintainer Isabella Gollini <igollini.stats@gmail.com>

Repository CRAN

Date/Publication 2015-07-08 14:48:30

R topics documented:

`tailloss-package` *Evaluate the Probability in the Upper Tail of the Aggregate Loss Distribution*

Description

Evaluate the probability in the upper tail of the aggregate loss distribution using different methods: Panjer recursion, Monte Carlo simulations, Markov bound, Cantelli bound, Moment bound, and Chernoff bound.

Details

The package `tailloss` contains functions to estimate the exceedance probability curve of the aggregated losses. There are two 'exact' approaches: Panjer recursion and Monte Carlo simulations, and four approaches producing upper bounds: the Markov bound, the Cantelli bound, the Moment bound, and the Chernoff bound. The upper bounds are useful and effective when the number of events in the catalogue is large, and there is interest in estimating the exceedance probabilities of exceptionally high losses.

Author(s)

Isabella Gollini <isabella.gollini@bristol.ac.uk>, and Jonathan Rougier.

This work was supported by the Natural Environment Research Council [Consortium on Risk in the Environment: Diagnostics, Integration, Benchmarking, Learning and Elicitation (CREDIBLE); grant number NE/J017450/1]

References

Gollini, I., and Rougier, J. C. (2015), "Rapidly bounding the exceedance probabilities of high aggregate losses", http://arxiv.org/abs/1507.01853.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
s <- seq(1,40)
EPC <- matrix(NA, length(s), 6)
colnames(EPC) <- c("Panjer", "MonteCarlo", "Markov",
 "Cantelli", "Moment", "Chernoff")
EPC[, 1] <- fPanjer(USh.m, s = s)[, 2]
EPC[, 2] <- fMonteCarlo(USh.m, s = s)[, 2]
EPC[, 3] <- fMarkov(USh.m, s = s)[, 2]
EPC[, 4] <- fCantelli(USh.m, s = s)[, 2]
EPC[, 5] <- fMoment(USh.m, s = s)[, 2]
EPC[, 6] <- fChernoff(USh.m, s = s)[, 2]
```

```
matplot(s, EPC, type = "l", lwd = 2, xlab = "s", ylim = c(0, 1), lty = 1:6,
  ylab = expression(plain(Pr)(S>=s)), main = "Exceedance Probability Curve")
zoombox(s, EPC, x0 = c(30, 40), y0 = c(0, .1), y1 = c(.3, .6), type = "l",
  lwd = 2, lty = 1:6)
legend("topright", legend = colnames(EPC), lty = 1:6, col = 1:6, lwd = 2)

EPCcap <- matrix(NA, length(s), 6)
colnames(EPCcap) <- c("Panjer", "MonteCarlo", "Markov",
 "Cantelli", "Moment", "Chernoff")
EPCcap[, 1] <- fPanjer(USh.m, s = s, theta = 2, cap = 5)[, 2]
EPCcap[, 2] <- fMonteCarlo(USh.m, s = s, theta = 2, cap = 5)[, 2]
EPCcap[, 3] <- fMarkov(USh.m, s = s, theta = 2, cap = 5)[, 2]
EPCcap[, 4] <- fCantelli(USh.m, s = s, theta = 2, cap = 5)[, 2]
EPCcap[, 5] <- fMoment(USh.m, s = s, theta = 2, cap = 5)[, 2]
EPCcap[, 6] <- fChernoff(USh.m, s = s, theta = 2, cap =  5)[, 2]
matplot(s, EPCcap, type = "l", lwd = 2, xlab = "s", ylim = c(0, 1), lty = 1:6,
  ylab = expression(plain(Pr)(S>=s)), main = "Exceedance Probability Curve")
zoombox(s, EPCcap, x0 = c(30, 40), y0 = c(0, .1), y1 = c(.3, .6), type = "l",
  lwd = 2, lty = 1:6)
legend("topright", legend = colnames(EPC), lty = 1:6, col = 1:6, lwd = 2)
```

`compressELT` *Compress the event loss table*

Description

Function to merge losses of the same amount adding up their corresponding occurrence rates, and to round the losses to the 10^`digits` integer value.

Usage

```
compressELT(ELT, digits = 0)
```

Arguments

`ELT`	Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
`digits`	Integer. It specifies the rounding of the losses to the 10^`digits` integer value of the event loss table. `digits` < 0 decreases the precision of the calculation, but considerably decreases the time to perform it. If `digits` = 0 it only merges the losses of the same amount adding up their corresponding rates. The default value is `digits` = 0.

Value

Data frame containg two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.

Examples

```
data(UShurricane)

# Compress the table to thousands of dollars

USh.k <- compressELT(ELT(UShurricane), digits = -3)
summary(USh.k)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
summary(USh.m)
```

ELT *Event Loss Table*

Description

Function to create an ELT object

Usage

```
ELT(X = NULL, Rate = NULL, Loss = NULL, ID = NULL)
```

Arguments

`X`	Data frame containing at least two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
`Rate`	Positive numeric vector of arrival rates
`Loss`	Positive numeric vector of losses
`ID`	Vector event ID.

Value

An object ELT, a data frame with 3 columns. The column `ID` contains the ID of each event. The column `Rate` contains the arrival rates of a single occurrence of event. The column `Loss` contains the expected losses from each single occurrence of event.

See Also

`data.frame`

Examples

```
rate <- c(.1, .02, .05)
loss <- c(2, 5, 7)

ELT(Rate = rate, Loss = loss)
# Same as
rl <- data.frame(Rate = rate, Loss = loss)
ELT(rl)
```

---

fCantelli *Cantelli Bound.*

Description

Function to bound the total losses via the Cantelli inequality.

Usage

```
fCantelli(ELT, s, t = 1, theta = 0, cap = Inf)
```

Arguments

ELT	Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
s	Scalar or numeric vector containing the total losses of interest.
t	Scalar representing the time period of interest. The default value is `t` = 1.
theta	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$`theta`. The default value is `theta` = 0: the loss associated to an event is considered as a constant.
cap	Scalar representing the level of truncation of the Gamma distribution, i.e. the maximum possible loss caused by a single event. The default value is `cap` = Inf.

Details

Cantelli's inequality states:

$$\Pr(S \geq s) \leq \frac{\sigma^2}{\sigma^2 + (s-\mu)^2} \quad \text{for } s \geq \mu,$$

where $\mu = E[S]$ and $\sigma^2 = Var[S] < \infty$ are the mean and the variance of the distribution of S.

Value

A numeric matrix, containing the pre-specified losses `s` in the first column and the upper bound for the exceedance probabilities in the second column.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
EPC.Cantelli  <- fCantelli(USh.m, s = 1:40)
plot(EPC.Cantelli, type = "l", ylim = c(0, 1))
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x
EPC.Cantelli.Gamma  <- fCantelli(USh.m, s = 1:40, theta = 2, cap = 25)
EPC.Cantelli.Gamma
plot(EPC.Cantelli.Gamma, type = "l")
# Compare the two results:
plot(EPC.Cantelli, type = "l", main = "Exceedance Probability Curve", ylim = c(0, 1))
lines(EPC.Cantelli.Gamma, col = 2, lty = 2)
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

---

fChernoff *Chernoff Bound.*

Description

Function to bound the total losses via the Chernoff inequality.

Usage

```
fChernoff(ELT, s, t = 1, theta = 0, cap = Inf, nk = 1001,
  verbose = FALSE)
```

Arguments

ELT	Data frame containing two numeric columns. The column Loss contains the expected losses from each single occurrence of event. The column Rate contains the arrival rates of a single occurrence of event.
s	Scalar or numeric vector containing the total losses of interest.
t	Scalar representing the time period of interest. The default value is t = 1.
theta	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$theta. The default value is theta = 0: the loss associated to an event is considered as a constant.
cap	Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is cap = Inf.
nk	Number of optimisation points.
verbose	Logical. If TRUE attaches the minimising index. The default is verbose = FALSE.

Details

Chernoff's inequality states:

$$\Pr(S \geq s) \leq \inf_{k>0} e^{-ks} M_S(k)$$

where $M_S(k)$ is the Moment Generating Function (MGF) of the total loss S. The `fChernoff` function optimises the bound over a fixed set of `nk` discrete values.

Value

A numeric matrix, containing the pre-specified losses `s` in the first column and the upper bound for the exceedance probabilities in the second column.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
EPC.Chernoff <- fChernoff(USh.m, s = 1:40)
EPC.Chernoff
plot(EPC.Chernoff, type = "l", ylim = c(0, 1))
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x
EPC.Chernoff.Gamma <- fChernoff(USh.m, s = 1:40, theta = 2, cap = 5)
EPC.Chernoff.Gamma
plot(EPC.Chernoff.Gamma, type = "l", ylim = c(0, 1))
# Compare the two results:
plot(EPC.Chernoff, type = "l", main = "Exceedance Probability Curve", ylim = c(0, 1))
lines(EPC.Chernoff.Gamma, col = 2, lty = 2)
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

`fMarkov` *Markov Bound.*

Description

Function to bound the total losses via the Markov inequality.

Usage

```
fMarkov(ELT, s, t = 1, theta = 0, cap = Inf)
```

Arguments

- `ELT` — Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
- `s` — Scalar or numeric vector containing the total losses of interest.
- `t` — Scalar representing the time period of interest. The default value is `t` = 1.
- `theta` — Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$`theta`. The default value is `theta` = 0: the loss associated to an event is considered as a constant.
- `cap` — Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is `cap` = Inf.

Details

Cantelli's inequality states:

$$\Pr(S \geq s) \leq \frac{E[S]}{s}$$

Value

A numeric matrix, containing the pre-specified losses `s` in the first column and the upper bound for the exceedance probabilities in the second column.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
EPC.Markov <- fMarkov(USh.m, s = 1:40)
plot(EPC.Markov, type = "l", ylim = c(0, 1))
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x
EPC.Markov.Gamma <- fMarkov(USh.m, s = 1:40, theta = 2, cap = 5)
EPC.Markov.Gamma
plot(EPC.Markov.Gamma, type = "l", ylim = c(0, 1))
# Compare the two results:
plot(EPC.Markov, type = "l", main = "Exceedance Probability Curve", ylim = c(0,1))
lines(EPC.Markov.Gamma, col = 2, lty = 2)
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

`fMoment` *Moment Bound.*

Description

Function to bound the total losses via the Moment inequality.

Usage

```
fMoment(ELT, s, t = 1, theta = 0, cap = Inf, verbose = FALSE)
```

Arguments

`ELT`	Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
`s`	Scalar or numeric vector containing the total losses of interest.
`t`	Scalar representing the time period of interest. The default value is `t` = 1.
`theta`	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$`theta`. The default value is `theta` = 0: the loss associated to an event is considered as a constant.
`cap`	Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is `cap` = Inf.
`verbose`	Logical. If `TRUE` attaches the minimising index. The default is `verbose` = FALSE.

Details

Moment inequality states:

$$\Pr(S \geq s) \leq \min_{k=1,2...} \frac{E(S^k)}{s^k}$$

where $E(S^k)$ is the k-th moment of the total loss S distribution.

Value

A numeric matrix, containing the pre-specified losses `s` in the first column and the upper bound for the exceedance probabilities in the second column.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
EPC.Moment <- fMoment(USh.m, s = 1:40)
```

```
EPC.Moment
plot(EPC.Moment, type = "l", ylim = c(0, 1))
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x
EPC.Moment.Gamma <- fMoment(USh.m, s = 1:40, theta = 2, cap = 5)
EPC.Moment.Gamma
plot(EPC.Moment.Gamma, type = "l", ylim = c(0, 1))
# Compare the two results:
plot(EPC.Moment, type = "l", main = "Exceedance Probability Curve", ylim = c(0, 1))
lines(EPC.Moment.Gamma, col = 2, lty = 2)
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

fMonteCarlo *Monte Carlo Simulations.*

Description

Function to estimate the total losses via the Monte Carlo simulations.

Usage

```
fMonteCarlo(ELT, s, t = 1, theta = 0, cap = Inf, nsim = 10000,
  verbose = FALSE)
```

Arguments

`ELT`	Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
`s`	Scalar or numeric vector containing the total losses of interest.
`t`	Scalar representing the time period of interest. The default value is `t` = 1.
`theta`	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$`theta`. The default value is `theta` = 0: the loss associated to an event is considered as a constant.
`cap`	Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is `cap` = Inf.
`nsim`	Integer representing the number of Monte Carlo simulations. The default value is `nsim` = 10e3.
`verbose`	Logical, if `TRUE` returns 95% CB and raw sample. The default is `verbose` = `FALSE`.

Value

If verbose = FALSE the function returns a numeric matrix, containing in the first column the pre-specified losses s, and the estimated exceedance probabilities in the second column. If verbose = TRUE the function returns a numeric matrix containing four columns. The first column contains the losses s, the second column contains the estimated exceedance probabilities, the other columns contain the 95% confidence bands. The attributes of this matrix are a vector simS containing the simulated losses.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
EPC.MonteCarlo <- fMonteCarlo(USh.m, s = 1:40, verbose = TRUE)
EPC.MonteCarlo
par(mfrow = c(1, 2))
plot(EPC.MonteCarlo[, 1:2], type = "l", ylim = c(0, 1))
matlines(EPC.MonteCarlo[, -2], ylim = c(0, 1), lty = 2, col = 1)
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x and cap = 5m
EPC.MonteCarlo.Gamma <- fMonteCarlo(USh.m, s = 1:40, theta = 2, cap = 5, verbose = TRUE)
EPC.MonteCarlo.Gamma
plot(EPC.MonteCarlo.Gamma[, 1:2], type = "l", ylim = c(0, 1))
matlines(EPC.MonteCarlo.Gamma[, -2], ylim = c(0,1), lty = 2, col = 1)
# Compare the two results:
par(mfrow = c(1, 1))
plot(EPC.MonteCarlo[, 1:2], type = "l", main = "Exceedance Probability Curve",
ylim = c(0, 1))
lines(EPC.MonteCarlo.Gamma[, 1:2], col = 2, lty = 2)
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

---

fPanjer *Panjer Recursion.*

Description

Function to calculate the total losses via the Panjer recursion.

Usage

```
fPanjer(ELT, s, t = 1, theta = 0, cap = Inf, nq = 10, verbose = FALSE)
```

Arguments

ELT	Data frame containing two numeric columns. The column Loss contains the expected losses from each single occurrence of event. The column Rate contains the arrival rates of a single occurrence of event.
s	Scalar or numeric vector containing the total losses of interest.
t	Scalar representing the time period of interest. The default value is t = 1.
theta	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$theta. The default value is theta = 0: the loss associated to an event is considered as a constant.
cap	Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is cap = Inf.
nq	Scalar, number of quantiles added when theta > 0
verbose	A logical, if TRUE gives the entire distribution up to the maximum value of s. If FALSE gives only the results for the specified values of s. The default is verbose = FALSE.

Value

A numeric matrix containing the pre-specified losses s in the first column and the exceedance probabilities in the second column.

References

Panjer, H.H. (1980), 'The aggregate claims distribution and stop-loss reinsurance,' *Transactions of the Society of Actuaries*, 32, 523-545.

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)

EPC.Panjer <- fPanjer(USh.m, s = 1:40, verbose = TRUE)
EPC.Panjer
plot(EPC.Panjer, type = "l", ylim = c(0,1))
# Assuming the losses follow a Gamma with E[X] = x, and Var[X] = 2 * x and cap = 5m

EPC.Panjer.Gamma <- fPanjer(USh.m, s = 1:40, theta = 2, cap = 5, verbose = TRUE)
EPC.Panjer.Gamma
plot(EPC.Panjer.Gamma, type = "l", ylim = c(0,1))

# Compare the two results:

plot(EPC.Panjer, type = "l", main = 'Exceedance Probability Curve',
ylim = c(0, 1))
lines(EPC.Panjer.Gamma, col = 2, lty = 2)
```

```
legend("topright", c("Dirac Delta", expression(paste("Gamma(",
alpha[i] == 1 / theta^2, ", ", beta[i] ==1 / (x[i] * theta^2), ")", " cap =", 5))),
lwd = 2, lty = 1:2, col = 1:2)
```

---

summary.ELT *Summary statistics for class ELT.*

Description

Summary statistics for class ELT.

Usage

```
## S3 method for class 'ELT'
summary(object, theta = 0, cap = Inf, t = 1, ...)
```

Arguments

`object`	An object of class `ELT`. Data frame containing two numeric columns. The column `Loss` contains the expected losses from each single occurrence of event. The column `Rate` contains the arrival rates of a single occurrence of event.
`theta`	Scalar containing information about the variance of the Gamma distribution: $sd[X] = x*$`theta`. The default value is `theta` = 0: the loss associated to an event is considered as a constant.
`cap`	Scalar representing the financial cap on losses for a single event, i.e. the maximum possible loss caused by a single event. The default value is `cap` = Inf.
`t`	Scalar representing the time period of interest. The default value is `t` = 1.
`...`	additional arguments affecting the summary produced.

Value

A list containing the data summary, and the means and the standard deviations of N, Y, and S.

Examples

```
data(UShurricane)
summary(ELT(UShurricane))
```

`UShurricane` *US hurricane data*

Description

US hurricane data provided by Peter Taylor and Dickie Whitaker.

Format

Data frame with 32060 rows and 3 columns

Details

- EventID. ID of 32060 events.
- Rate. Annual rate of occurrence.
- Loss. Loss associated to each event measured in $.

`zoombox` *Function for zooming onto a matplot(x, y, ...).*

Description

Function for zooming onto a `matplot(x, y, ...)`.

Usage

```
zoombox(x, y, x0, y0 = c(0, 0.05), y1 = c(0.1, 0.6), ...)
```

Arguments

`x,y`	Vectors or matrices of data for plotting. The number of rows should match. If one of them are missing, the other is taken as y and an x vector of 1:n is used. Missing values (NAs) are allowed.
`x0`	range of x to zoom on.
`y0`	range of y to zoom on. The default value is `y0 = c(0,0.05)`
`y1`	range of y where to put the zoomed area. The default value is `y1 = c(0.1,0.6)`
`...`	Arguments to be passed to methods, such as graphical parameters (see `par`).

See Also

`matplot`, `plot`

Examples

```
data(UShurricane)

# Compress the table to millions of dollars

USh.m <- compressELT(ELT(UShurricane), digits = -6)
s <- seq(1,40)
EPC <- matrix(NA, length(s), 6)
colnames(EPC) <- c("Panjer", "MonteCarlo", "Markov",
 "Cantelli", "Moment", "Chernoff")
EPC[, 1] <- fPanjer(USh.m, s = s)[, 2]
EPC[, 2] <- fMonteCarlo(USh.m, s = s)[, 2]
EPC[, 3] <- fMarkov(USh.m, s = s)[, 2]
EPC[, 4] <- fCantelli(USh.m, s = s)[, 2]
EPC[, 5] <- fMoment(USh.m, s = s)[, 2]
EPC[, 6] <- fChernoff(USh.m, s = s)[, 2]
matplot(s, EPC, type = "l", lwd = 2, xlab = "s", ylim = c(0, 1), lty = 1:6,
  ylab = expression(plain(Pr)(S>=s)), main = "Exceedance Probability Curve")
zoombox(s, EPC, x0 = c(30, 40), y0 = c(0, .1), y1 = c(.3, .6), type = "l", lwd = 2, lty = 1:6)
legend("topright", legend = colnames(EPC), lty = 1:6, col = 1:6, lwd = 2)
```

Index