Package ‘mvtnorm’

January 28, 2011

Title Multivariate Normal and t Distributions

Version 0.9-96

Date 2011-01-28

Author Alan Genz, Frank Bretz, Tetsuhisa Miwa, Xuefei Mi, Friedrich Leisch, Fabian Scheipl, Bjoern Bornkamp, Torsten Hothorn

Maintainer Torsten Hothorn <Torsten.Hothorn@R-project.org>

Description
Computes multivariate normal and t probabilities,quantiles, random deviates and densities.

Imports stats

Depends R(>= 1.9.0)

License GPL-2

Repository CRAN

Date/Publication 2011-01-28 16:01:36

R topics documented:

`algorithms` *Choice of Algorithm and Hyper Parameters*

Description

Choose between three algorithms for evaluating normal distributions and define hyper parameters.

Usage

```
GenzBretz(maxpts = 25000, abseps = 0.001, releps = 0)
Miwa(steps = 128)
TVPACK(abseps = 1e-6)
```

Arguments

`maxpts`	maximum number of function values as integer.
`abseps`	absolute error tolerance as double.
`releps`	relative error tolerance as double.
`steps`	number of grid points to be evaluated.

Details

There are three algorithms available for evaluating normal probabilities: The default is the randomized Quasi-Monte-Carlo procedure by Genz (1992, 1993) and Genz and Bretz (2002) applicable to arbitrary covariance structures and dimensions up to 1000.

For smaller dimensions (up to 20) and non-singular covariance matrices, the algorithm by Miwa et al. (2003) can be used as well.

For two- and three-dimensional problems and semi-infinite integration region, `TVPACK` implements an interface to the methods described by Genz (2004).

Value

An object of class `GenzBretz` or `Miwa` defining hyper parameters.

References

Genz, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, **1**, 141–150.

Genz, A. (1993). Comparison of methods for the computation of multivariate normal probabilities. *Computing Science and Statistics*, **25**, 400–405.

Genz, A. and Bretz, F. (2002), Methods for the computation of multivariate t-probabilities. *Journal of Computational and Graphical Statistics*, **11**, 950–971.

Genz, A. (2004), Numerical computation of rectangular bivariate and trivariate normal and t-probabilities, *Statistics and Computing*, **14**, 251–260.

Genz, A. and Bretz, F. (2009), *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics, Vol. 195. Springer-Verlag, Heidelberg.

Miwa, A., Hayter J. and Kuriki, S. (2003). The evaluation of general non-centred orthant probabilities. *Journal of the Royal Statistical Society*, Ser. B, 65, 223–234.

`Mvnorm` *Multivariate Normal Density and Random Deviates*

Description

These functions provide the density function and a random number generator for the multivariate normal distribution with mean equal to `mean` and covariance matrix `sigma`.

Usage

```
dmvnorm(x, mean, sigma, log=FALSE)
rmvnorm(n, mean = rep(0, nrow(sigma)), sigma = diag(length(mean)),
        method=c("eigen", "svd", "chol"))
```

Arguments

`x`	Vector or matrix of quantiles. If `x` is a matrix, each row is taken to be a quantile.
`n`	Number of observations.
`mean`	Mean vector, default is `rep(0, length = ncol(x))`.
`sigma`	Covariance matrix, default is `diag(ncol(x))`.
`log`	Logical; if `TRUE`, densities d are given as log(d).
`method`	Matrix decomposition used to determine the matrix root of `sigma`, possible methods are eigenvalue decomposition (`"eigen"`, default), singular value decomposition (`"svd"`), and Cholesky decomposition (`"chol"`).

Author(s)

Friedrich Leisch and Fabian Scheipl

See Also

`pmvnorm`, `rnorm`, `qmvnorm`

Examples

```
dmvnorm(x=c(0,0))
dmvnorm(x=c(0,0), mean=c(1,1))

sigma <- matrix(c(4,2,2,3), ncol=2)
x <- rmvnorm(n=500, mean=c(1,2), sigma=sigma)
colMeans(x)
var(x)

x <- rmvnorm(n=500, mean=c(1,2), sigma=sigma, method="chol")
colMeans(x)
var(x)

plot(x)
```

Mvt *The Multivariate t Distribution*

Description

These functions provide information about the multivariate t distribution with non-centrality parameter (or mode) `delta`, covariance matrix `sigma` and degrees of freedom `df`. `dmvt` gives the density and `rmvt` generates random deviates.

Usage

```
rmvt(n, sigma = diag(2), df = 1, delta = rep(0, nrow(sigma)),
     type = c("shifted", "Kshirsagar"))
dmvt(x, delta, sigma, df = 1, log = TRUE,
     type = "shifted")
```

Arguments

`x`	Vector or matrix of quantiles. If `x` is a matrix, each row is taken to be a quantile.
`n`	Number of observations.
`delta`	the vector of noncentrality parameters of length n, for `type = "shifted"` delta specifies the mode.
`sigma`	Covariance matrix, default is `diag(ncol(x))`.
`df`	degree of freedom as integer.
`log`	Logical; if `TRUE`, densities d are given as log(d).
`type`	type of the noncentral multivariate t distribution to be computed. `type = "Kshirsagar"` corresponds to formula (1.4) in Genz and Bretz (2009) (see also Chapter 5.1 in Kotz and Nadarajah (2004)). This is the noncentral t-distribution needed for calculating the power of multiple contrast tests under a normality assumption. `type = "shifted"` corresponds to the formula right before formula (1.4) in Genz and Bretz (2009) (see also formula (1.1) in Kotz and

Nadarajah (2004)). It is a location shifted version of the central t-distribution. This noncentral multivariate t distribution appears for example as the Bayesian posterior distribution for the regression coefficients in a linear regression. In the central case both types coincide. Note that the defaults differ from the default in `pmvt` (for reasons of backward compatibility).

Details

For type = "shifted" the following density is implemented

$$c(1 + (x - \delta)' S^{-1} (x - \delta)/\nu)^{-(\nu+m)/2},$$

where

$$c = \Gamma((\nu + m)/2)/((\pi\nu)^{m/2}\Gamma(\nu/2)|S|^{1/2}),$$

here S is a positive definite symmetric matrix (which might be the correlation or the covariance matrix), $delta$ is the non-centrality vector and ν are the degrees of freedom.

See Also

`pmvt` and `qmvt`

Examples

```
dmvt(x=c(0,0), sigma = diag(2))
x <- rmvt(n=100, sigma = diag(2), df = 3)
plot(x)
```

`pmvnorm` *Multivariate Normal Distribution*

Description

Computes the distribution function of the multivariate normal distribution for arbitrary limits and correlation matrices.

Usage

```
pmvnorm(lower=-Inf, upper=Inf, mean=rep(0, length(lower)),
        corr=NULL, sigma=NULL, algorithm = GenzBretz(), ...)
```

Arguments

`lower`	the vector of lower limits of length n.
`upper`	the vector of upper limits of length n.
`mean`	the mean vector of length n.
`corr`	the correlation matrix of dimension n.
`sigma`	the covariance matrix of dimension n. Either `corr` or `sigma` can be specified. If `sigma` is given, the problem is standardized. If neither `corr` nor `sigma` is given, the identity matrix is used for `sigma`.
`algorithm`	an object of class `GenzBretz`, `Miwa` or `TVPACK` specifying both the algorithm to be used as well as the associated hyper parameters.
`...`	additional parameters (currently given to `GenzBretz` for backward compatibility issues).

Details

This program involves the computation of multivariate normal probabilities with arbitrary correlation matrices. It involves both the computation of singular and nonsingular probabilities. The implemented methodology is described in Genz (1992, 1993) (for algorithm GenzBretz), in Miwa et al. (2003) for algorithm Miwa (useful up to dimension 20) and Genz (2004) for the TVPACK algorithm (which covers 2- and 3-dimensional problems for semi-infinite integration regions).

Note that both `-Inf` and `+Inf` may be specified in `lower` and `upper`. For more details see `pmvt`.

The multivariate normal case is treated as a special case of `pmvt` with `df=0` and univariate problems are passed to `pnorm`.

The multivariate normal density and random deviates are available using `dmvnorm` and `rmvnorm`.

Value

The evaluated distribution function is returned with attributes

`error`	estimated absolute error and
`msg`	status messages.

Source

`http://www.sci.wsu.edu/math/faculty/genz/homepage`

References

Genz, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, **1**, 141–150.

Genz, A. (1993). Comparison of methods for the computation of multivariate normal probabilities. *Computing Science and Statistics*, **25**, 400–405.

Genz, A. (2004), Numerical computation of rectangular bivariate and trivariate normal and t-probabilities, *Statistics and Computing*, **14**, 251–260.

Genz, A. and Bretz, F. (2009), *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics, Vol. 195. Springer-Verlag, Heidelberg.

Miwa, A., Hayter J. and Kuriki, S. (2003). The evaluation of general non-centred orthant probabilities. *Journal of the Royal Statistical Society*, Ser. B, 65, 223–234.

See Also

qmvnorm

Examples

```
n <- 5
mean <- rep(0, 5)
lower <- rep(-1, 5)
upper <- rep(3, 5)
corr <- diag(5)
corr[lower.tri(corr)] <- 0.5
corr[upper.tri(corr)] <- 0.5
prob <- pmvnorm(lower, upper, mean, corr)
print(prob)

stopifnot(pmvnorm(lower=-Inf, upper=3, mean=0, sigma=1) == pnorm(3))

a <- pmvnorm(lower=-Inf,upper=c(.3,.5),mean=c(2,4),diag(2))

stopifnot(round(a,16) == round(prod(pnorm(c(.3,.5),c(2,4))),16))

a <- pmvnorm(lower=-Inf,upper=c(.3,.5,1),mean=c(2,4,1),diag(3))

stopifnot(round(a,16) == round(prod(pnorm(c(.3,.5,1),c(2,4,1))),16))

# Example from R News paper (original by Genz, 1992):

m <- 3
sigma <- diag(3)
sigma[2,1] <- 3/5
sigma[3,1] <- 1/3
sigma[3,2] <- 11/15
pmvnorm(lower=rep(-Inf, m), upper=c(1,4,2), mean=rep(0, m), corr=sigma)

# Correlation and Covariance

a <- pmvnorm(lower=-Inf, upper=c(2,2), sigma = diag(2)*2)
b <- pmvnorm(lower=-Inf, upper=c(2,2)/sqrt(2), corr=diag(2))
stopifnot(all.equal(round(a,5) , round(b, 5)))
```

pmvt *Multivariate t Distribution*

Description

Computes the the distribution function of the multivariate t distribution for arbitrary limits, degrees of freedom and correlation matrices based on algorithms by Genz and Bretz.

Usage

```
pmvt(lower=-Inf, upper=Inf, delta=rep(0, length(lower)),
     df=1, corr=NULL, sigma=NULL, algorithm = GenzBretz(),
     type = c("Kshirsagar", "shifted"), ...)
```

Arguments

`lower`	the vector of lower limits of length n.
`upper`	the vector of upper limits of length n.
`delta`	the vector of noncentrality parameters of length n, for `type = "shifted"` delta specifies the mode.
`df`	degree of freedom as integer. Normal probabilities are computed for `df=0`.
`corr`	the correlation matrix of dimension n.
`sigma`	the covariance matrix of dimension n. Either `corr` or `sigma` can be specified. If `sigma` is given, the problem is standardized. If neither `corr` nor `sigma` is given, the identity matrix is used for `sigma`.
`algorithm`	an object of class `GenzBretz` or `TVPACK` defining the hyper parameters of this algorithm.
`type`	type of the noncentral multivariate t distribution to be computed. `type = "Kshirsagar"` corresponds to formula (1.4) in Genz and Bretz (2009) (see also Chapter 5.1 in Kotz and Nadarajah (2004)). This is the noncentral t-distribution needed for calculating the power of multiple contrast tests under a normality assumption. `type = "shifted"` corresponds to the formula right before formula (1.4) in Genz and Bretz (2009) (see also formula (1.1) in Kotz and Nadarajah (2004)). It is a location shifted version of the central t-distribution. This noncentral multivariate t distribution appears for example as the Bayesian posterior distribution for the regression coefficients in a linear regression. In the central case both types coincide.
`...`	additional parameters (currently given to `GenzBretz` for backward compatibility issues).

Details

This program involves the computation of central and noncentral multivariate t-probabilities with arbitrary correlation matrices. It involves both the computation of singular and nonsingular probabilities. The methodology is based on randomized quasi Monte Carlo methods and described in Genz and Bretz (1999, 2002).

For 2- and 3-dimensional problems one can also use the TVPACK routines described by Genz (2004), which only handles semi-infinite integration regions (and for `type = "Kshirsagar"` only central problems).

For `type = "Kshirsagar"` and a given correlation matrix `corr`, for short A, say, (which has to be positive semi-definite) and degrees of freedom ν the following values are numerically evaluated

$$I = 2^{1-\nu/2}/\Gamma(\nu/2) \int_0^\infty s^{\nu-1} \exp(-s^2/2) \Phi(s \cdot lower/\sqrt{\nu} - \delta, s \cdot upper/\sqrt{\nu} - \delta)\, ds$$

where

$$\Phi(a,b) = (det(A)(2\pi)^m)^{-1/2} \int_a^b \exp(-x' Ax/2)\, dx$$

is the multivariate normal distribution and m is the number of rows of A.

For `type = "shifted"`, a positive definite symmetric matrix S (which might be the correlation or the covariance matrix), non-centrality vector δ and degrees of freedom ν the following integral is evaluated:

$$c \int_{lower_1}^{upper_1} \dots \int_{lower_m}^{upper_m} (1 + (x-\delta)' S^{-1} (x-\delta)/\nu)^{-(\nu+m)/2}\, dx_1 \dots dx_m,$$

where

$$c = \Gamma((\nu+m)/2)/((\pi\nu)^{m/2}\Gamma(\nu/2)|S|^{1/2}),$$

and m is the number of rows of S.

Note that both `-Inf` and `+Inf` may be specified in the lower and upper integral limits in order to compute one-sided probabilities.

Univariate problems are passed to `pt`. If `df = 0`, normal probabilities are returned.

Value

The evaluated distribution function is returned with attributes

`error`	estimated absolute error and
`msg`	status messages.

Source

http://www.sci.wsu.edu/math/faculty/genz/homepage

References

Genz, A. and Bretz, F. (1999), Numerical computation of multivariate t-probabilities with application to power calculation of multiple contrasts. *Journal of Statistical Computation and Simulation*, **63**, 361–378.

Genz, A. and Bretz, F. (2002), Methods for the computation of multivariate t-probabilities. *Journal of Computational and Graphical Statistics*, **11**, 950–971.

Genz, A. (2004), Numerical computation of rectangular bivariate and trivariate normal and t-probabilities, *Statistics and Computing*, **14**, 251–260.

Genz, A. and Bretz, F. (2009), *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics, Vol. 195. Springer-Verlag, Heidelberg.

S. Kotz and S. Nadarajah (2004), *Multivariate t Distributions and Their Applications*. Cambridge University Press. Cambridge.

Edwards D. and Berry, Jack J. (1987), The efficiency of simulation-based multiple comparisons. *Biometrics*, **43**, 913–928.

See Also

qmvt

Examples

```
n <- 5
lower <- -1
upper <- 3
df <- 4
corr <- diag(5)
corr[lower.tri(corr)] <- 0.5
delta <- rep(0, 5)
prob <- pmvt(lower=lower, upper=upper, delta=delta, df=df, corr=corr)
print(prob)

pmvt(lower=-Inf, upper=3, df = 3, sigma = 1) == pt(3, 3)

# Example from R News paper (original by Edwards and Berry, 1987)

n <- c(26, 24, 20, 33, 32)
V <- diag(1/n)
df <- 130
C <- c(1,1,1,0,0,-1,0,0,1,0,0,-1,0,0,1,0,0,0,-1,-1,0,0,-1,0,0)
C <- matrix(C, ncol=5)
### covariance matrix
cv <- C %*% V %*% t(C)
### correlation matrix
dv <- t(1/sqrt(diag(cv)))
cr <- cv * (t(dv) %*% dv)
delta <- rep(0,5)

myfct <- function(q, alpha) {
```

```
  lower <- rep(-q, ncol(cv))
  upper <- rep(q, ncol(cv))
  pmvt(lower=lower, upper=upper, delta=delta, df=df,
       corr=cr, abseps=0.0001) - alpha
}

round(uniroot(myfct, lower=1, upper=5, alpha=0.95)$root, 3)

# compare pmvt and pmvnorm for large df:

a <- pmvnorm(lower=-Inf, upper=1, mean=rep(0, 5), corr=diag(5))
b <- pmvt(lower=-Inf, upper=1, delta=rep(0, 5), df=rep(300,5),
          corr=diag(5))
a
b

stopifnot(round(a, 2) == round(b, 2))

# correlation and covariance matrix

a <- pmvt(lower=-Inf, upper=2, delta=rep(0,5), df=3,
          sigma = diag(5)*2)
b <- pmvt(lower=-Inf, upper=2/sqrt(2), delta=rep(0,5),
          df=3, corr=diag(5))
attributes(a) <- NULL
attributes(b) <- NULL
a
b
stopifnot(all.equal(round(a,3) , round(b, 3)))

a <- pmvt(0, 1,df=10)
attributes(a) <- NULL
b <- pt(1, df=10) - pt(0, df=10)
stopifnot(all.equal(round(a,10) , round(b, 10)))
```

`qmvnorm` *Quantiles of the Multivariate Normal Distribution*

Description

Computes the equicoordinate quantile function of the multivariate normal distribution for arbitrary correlation matrices based on inversion of pmvnorm.

Usage

```
qmvnorm(p, interval = NULL, tail = c("lower.tail",
        "upper.tail", "both.tails"), mean = 0, corr = NULL,
        sigma = NULL, algorithm = GenzBretz(), ...)
```

Arguments

`p`	probability.
`interval`	optional, a vector containing the end-points of the interval to be searched by `uniroot`.
`tail`	specifies which quantiles should be computed. `lower.tail` gives the quantile x for which $P[X \leq x] = p$, `upper.tail` gives x with $P[X > x] = p$ and `both.tails` leads to x with $P[-x \leq X \leq x] = p$.
`mean`	the mean vector of length n.
`corr`	the correlation matrix of dimension n.
`sigma`	the covariance matrix of dimension n. Either `corr` or `sigma` can be specified. If `sigma` is given, the problem is standardized. If neither `corr` nor `sigma` is given, the identity matrix is used for `sigma`.
`algorithm`	an object of class `GenzBretz`, `Miwa` or `TVPACK` specifying both the algorithm to be used as well as the associated hyper parameters.
`...`	additional parameters to be passed to `GenzBretz`.

Details

Only equicoordinate quantiles are computed, i.e., the quantiles in each dimension coincide. Currently, the distribution function is inverted by using the `uniroot` function which may result in limited accuracy of the quantiles.

Value

A list with four components: `quantile` and `f.quantile` give the location of the quantile and the value of the function evaluated at that point. `iter` and `estim.prec` give the number of iterations used and an approximate estimated precision from `uniroot`.

See Also

`pmvnorm`, `qmvt`

Examples

```
qmvnorm(0.95, sigma = diag(2), tail = "both")
```

---

qmvt — *Quantiles of the Multivariate t Distribution*

Description

Computes the equicoordinate quantile function of the multivariate t distribution for arbitrary correlation matrices based on inversion of qmvt.

Usage

```
qmvt(p, interval = NULL, tail = c("lower.tail",
     "upper.tail", "both.tails"), df = 1, delta = 0, corr = NULL,
     sigma = NULL, algorithm = GenzBretz(),
     type = c("Kshirsagar", "shifted"), ...)
```

Arguments

`p`	probability.
`interval`	optional, a vector containing the end-points of the interval to be searched by `uniroot`.
`tail`	specifies which quantiles should be computed. `lower.tail` gives the quantile x for which $P[X \leq x] = p$, `upper.tail` gives x with $P[X > x] = p$ and `both.tails` leads to x with $P[-x \leq X \leq x] = p$.
`delta`	the vector of noncentrality parameters of length n, for `type = "shifted"` delta specifies the mode.
`df`	degree of freedom as integer. Normal quantiles are computed for `df = 0`.
`corr`	the correlation matrix of dimension n.
`sigma`	the covariance matrix of dimension n. Either `corr` or `sigma` can be specified. If `sigma` is given, the problem is standardized. If neither `corr` nor `sigma` is given, the identity matrix is used for `sigma`.
`algorithm`	an object of class `GenzBretz` or `TVPACK` defining the hyper parameters of this algorithm.
`type`	type of the noncentral multivariate t distribution to be computed. `type = "Kshirsagar"` corresponds to formula (1.4) in Genz and Bretz (2009) (see also Chapter 5.1 in Kotz and Nadarajah (2004)) and `type = "shifted"` corresponds to the formula before formula (1.4) in Genz and Bretz (2009) (see also formula (1.1) in Kotz and Nadarajah (2004)).
`...`	additional parameters to be passed to `GenzBretz`.

Details

Only equicoordinate quantiles are computed, i.e., the quantiles in each dimension coincide. Currently, the distribution function is inverted by using the `uniroot` function which may result in limited accuracy of the quantiles.

Value

A list with four components: `quantile` and `f.quantile` give the location of the quantile and the value of the function evaluated at that point. `iter` and `estim.prec` give the number of iterations used and an approximate estimated precision from `uniroot`.

See Also

`pmvnorm`, `qmvnorm`

Examples

```
qmvt(0.95, df = 16, tail = "both")
```

Index