Package ‘gaussratiovegind’

June 16, 2025

Title Distribution of Gaussian Ratios

Version 2.0.3

Maintainer Pierre Santagostini <pierre.santagostini@institut-agro.fr>

Description It is well known that the distribution of a Gaussian ratio does not follow a Gaussian distribution.
The lack of awareness among users of vegetation indices about this non-Gaussian nature could lead to incorrect statistical modeling and interpretation.
This package provides tools to accurately handle and analyse such ratios: density function, parameter estimation, simulation.
An example on the study of chlorophyll fluorescence can be found in A. El Ghaziri et al. (2023) <doi:10.3390/rs15020528>
and another method for parameter estimation is given in Bouhlel et al. (2023) <doi:10.23919/EUSIPCO58844.2023.10290111>.

License GPL (>= 3)

URL https://forge.inrae.fr/imhorphen/gaussratiovegind

BugReports https://forge.inrae.fr/imhorphen/gaussratiovegind/-/issues

Encoding UTF-8

RoxygenNote 7.3.2

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Depends R (>= 3.5)

LazyData true

NeedsCompilation no

Author Pierre Santagostini [aut, cre],
Angélina El Ghaziri [aut],
Nizar Bouhlel [aut],
David Rousseau [ctb]

Repository CRAN

Date/Publication 2025-06-16 17:00:03 UTC

Contents

`arabidopsis` *Statistics on Chlorophyll Fluorescence Parameters*

Description

Mean and standard deviation values on healthy and diseased tissues of chlorophyll fluorescence parameters F_0 (minimum fluorescence) and F_m (maximum fluorescence) for a dataset of *Arabidopsis thaliana* plants infected with fungal pathogen data; parameters of the distribution of the ratio $\frac{F_v}{F_m} = \frac{F_m - F_0}{F_m}$.

Usage

```
arabidopsis
```

Format

A data frame with 10 rows and 6 columns:

time times of the acquisition of chlorophyll fluorescence images

condition indicates if the plant was inoculated: `healthy` (inoculated with water) or `diseased` (inoculated with the pathogen)

mF0, sF0 Mean and standard deviation values of the chlorophyll parameter F_0

mFm, sFm Mean and standard deviation values of the chlorophyll parameter F_m

beta, rho, delta the β, ρ and δ_y parameters of the distribution of $\frac{F_v}{F_m} = \frac{F_m - F_0}{F_m}$ (distributed according to a normal ratio distribution, see Details)

Details

On each leaf picture, the F_0 and F_m values are normally distributed. Hence, $\frac{F_0}{F_m}$ is a ratio of two normal distributions.

Let μ_{F_0} and σ_{F_0} the mean and standard deviation of F_0 and μ_{F_m} and σ_{F_m} the mean and standard deviation of F_m. The parameters β, ρ and δ_y are given by:

$$\beta = \frac{\mu_{F_0}}{\mu_{F_m}}$$

$$\rho = \frac{\sigma_{F_m}}{\sigma_{F_0}}$$

$$\delta_y = \frac{\sigma_{F_m}}{\mu_{F_m}}$$

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

Pavicic, M., Overmyer, K., Rehman, A.u., Jones, P., Jacobson, D., Himanen, K. Image-Based Methods to Score Fungal Pathogen Symptom Progression and Severity in Excised Arabidopsis Leaves. Plants, 10, 158 (2021). doi:10.3390/plants10010158

`dnormratio` *Density Function of a Normal Ratio Distribution*

Description

Density of the ratio of two independent Gaussian distributions.

Usage

```
dnormratio(z, bet, rho, delta)
```

Arguments

`z`	length p numeric vector.
`bet, rho, delta`	numeric values. The parameters (β, ρ, δ_y) of the distribution, see Details.

Details

Let two independant random variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$.

If we denote $\beta = \frac{\mu_x}{\mu_y}$, $\rho = \frac{\sigma_y}{\sigma_x}$ and $\delta_y = \frac{\sigma_y}{\mu_y}$, the probability distribution function of the ratio $Z = \frac{X}{Y}$ is given by:

$$f_Z(z; \beta, \rho, \delta_y) = \frac{\rho}{\pi(1+\rho^2 z^2)} \left[\exp\left(-\frac{\rho^2\beta^2+1}{2\delta_y^2} \right) + \sqrt{\frac{\pi}{2}}\, q \operatorname{erf}\left(\frac{q}{\sqrt{2}} \right) \exp\left(-\frac{\rho^2(z-\beta)^2}{2\delta_y^2(1+\rho^2 z^2)} \right) \right]$$

with $q = \frac{1+\beta\rho^2 z}{\delta_y\sqrt{1+\rho^2 z^2}}$ and $\operatorname{erf}\left(\frac{q}{\sqrt{2}}\right) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{q}{\sqrt{2}}} \exp\left(-t^2\right) dt$

Another expression of this density, used by the estparnormratio() function, is:

$$f_Z(z;\beta,\rho,\delta_y) = \frac{\rho}{\pi(1+\rho^2 z^2)}\ \exp\left(-\frac{\rho^2\beta^2+1}{2\delta_y^2}\right) {}_1F_1\left(1,\frac{1}{2};\frac{1}{2\delta_y^2}\frac{(1+\beta\rho^2 z)^2}{1+\rho^2 z^2}\right)$$

where ${}_1F_1\left(a,b;x\right)$ is the confluent hypergeometric function (Kummer's function):

$${}_1F_1\left(a,b;x\right) = \sum_{n=0}^{+\infty}\frac{(a)_n}{(b)_n}\frac{x^n}{n!}$$

Value

Numeric: the value of density.

Author(s)

Pierre Santagostini, Angélina El Ghaziri, Nizar Bouhlel

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

Marsaglia, G. 2006. Ratios of Normal Variables. Journal of Statistical Software 16. doi:10.18637/jss.v016.i04

Díaz-Francés, E., Rubio, F.J., On the existence of a normal approximation to the distribution of the ratio of two independent normal random variables. Stat Papers 54, 309–323 (2013). doi:10.1007/s0036201204292

See Also

pnormratio(): probability distribution function.

rnormratio(): sample simulation.

estparnormratio(): parameter estimation.

Examples

```
# First example
beta1 <- 0.15
rho1 <- 5.75
delta1 <- 0.22
dnormratio(0, bet = beta1, rho = rho1, delta = delta1)
dnormratio(0.5, bet = beta1, rho = rho1, delta = delta1)
curve(dnormratio(x, bet = beta1, rho = rho1, delta = delta1), from = -0.1, to = 0.7)

# Second example
beta2 <- 2
rho2 <- 2
delta2 <- 2
dnormratio(0, bet = beta2, rho = rho2, delta = delta2)
```

```
dnormratio(0.5, bet = beta2, rho = rho2, delta = delta2)
curve(dnormratio(x, bet = beta2, rho = rho2, delta = delta2), from = -0.1, to = 0.7)
```

---

estparnormratio *Estimation of the Parameters of a Normal Ratio Distribution*

Description

Estimation of the parameters of a ratio $Z = \frac{X}{Y}$, X and Y being two independent random variables distributed according to Gaussian distributions, using the EM (estimation-maximization) algorithm or variational inference. Depending on the estimation method, the `estparnormatio` function calls `estparEM` (EM algorithm) or `estparVB` (variational Bayes).

Usage

```
estparnormratio(z, method = c("EM", "VB"), eps = 1e-06,
                       display = FALSE, mux0 = 1, sigmax0 = 1,
                       alphax0 = NULL, betax0 = NULL, muy0 = 1, sigmay0 = 1,
                       alphay0 = NULL, betay0 = NULL)

estparEM(z, eps = 1e-06,  display = FALSE, #plot = display,
                       mux0 = 1, sigmax0 = 1, muy0 = 1, sigmay0 = 1)

estparVB(z, eps = 1e-06, display = FALSE, mux0 = 1, sigmax0 = 1,
                       alphax0 = 1, betax0 = 1, muy0 = 1, sigmay0 = 1,
                       alphay0 = 1, betay0 = 1)

estparEM(
  z,
  eps = 1e-06,
  display = FALSE,
  mux0 = 1,
  sigmax0 = 1,
  muy0 = 1,
  sigmay0 = 1
)

estparVB(
  z,
  eps = 1e-06,
  display = FALSE,
  mux0 = 1,
  sigmax0 = 1,
  alphax0 = 1,
  betax0 = 1,
  muy0 = 1,
```

```
  sigmay0 = 1,
  alphay0 = 1,
  betay0 = 1
)
```

Arguments

- `z` — numeric.
- `method` — the method used to estimate the parameters of the distribution. It can be `"EM"` (expectation-maximization) or `"VB"` (Variational Bayes).
- `eps` — numeric. Precision for the estimation of the parameters (see Details).
- `display` — logical. When `TRUE` the successive values of the stop criterion (distance between successive values) is printed.
- `mux0`, `sigmax0`, `muy0`, `sigmay0` — initial values of the means and standard deviations of the X and Y variables. Default: `mux0 = 1`, `sigmax0 = 1`, `muy0 = 1`, `sigmay0 = 1`.
- `alphax0`, `betax0`, `alphay0`, `betay0` — initial values for the variational Bayes method. Omitted if `method="EM"`. If `method="VB"`, if omitted, they are set to 1.

Details

Let a random variable: $Z = \frac{X}{Y}$,

X and Y being normally distributed: $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$.

The density probability of Z is:

$$f_Z(z; \beta, \rho, \delta_y) = \frac{\rho}{\pi(1+\rho^2 z^2)} \exp\left(-\frac{\rho^2\beta^2+1}{2\delta_y^2}\right) {}_1F_1\left(1, \frac{1}{2}; \frac{1}{2\delta_y^2}\frac{(1+\beta\rho^2 z)^2}{1+\rho^2 z^2}\right)$$

with: $\beta = \frac{\mu_x}{\mu_y}$, $\rho = \frac{\sigma_y}{\sigma_x}$, $\delta_y = \frac{\sigma_y}{\mu_y}$.

and ${}_1F_1(a, b; x)$ is the confluent hypergeometric function (Kummer's function):

$${}_1F_1(a, b; x) = \sum_{n=0}^{+\infty} \frac{(a)_n}{(b)_n}\frac{x^n}{n!}$$

If `method = "EM"`, the means and standard deviations μ_x, σ_x, μ_y and σ_y are estimated with the EM algorithm, as presented in El Ghaziri et al. If `method = "VB"`, they are estimated with the variational Bayes method as presented in Bouhlel et al.

Then the parameters β, ρ, δ_y of the Z distribution are computed from these means and standard deviations.

The estimation of μ_x, σ_x, μ_y and σ_y uses an iterative algorithm. The precision for their estimation is given by the `eps` parameter.

The computation uses the `kummer` function.

If there are ties in the `z` vector, it generates a warning, as there should be no ties in data distributed among a continuous distribution.

Value

A list of 3 elements `beta`, `rho`, `delta`: the estimated parameters of the Z distribution $\hat{\beta}$, $\hat{\rho}$, $\hat{\delta}_y$, with three attributes `attr(, "epsilon")` (precision of the result), `attr(, "k")` (number of iterations) and `attr(, "method")` (estimation method).

Author(s)

Pierre Santagostini, Angélina El Ghaziri, Nizar Bouhlel

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

Bouhlel, N., Mercier, F., El Ghaziri, A., Rousseau, D., Parameter Estimation of the Normal Ratio Distribution with Variational Inference. 2023 31st European Signal Processing Conference (EUSIPCO), Helsinki, Finland, 2023, pp. 1823-1827. doi:10.23919/EUSIPCO58844.2023.10290111

See Also

`dnormratio()`: probability density of a normal ratio.

`rnormratio()`: sample simulation.

Examples

```
# First example
beta1 <- 0.15
rho1 <- 5.75
delta1 <- 0.22

set.seed(1234)
z1 <- rnormratio(800, bet = beta1, rho = rho1, delta = delta1)

# With the EM method:
estparnormratio(z1, method = "EM")

# With the variational method:
estparnormratio(z1, method = "VB")

# Second example
beta2 <- 0.24
rho2 <- 4.21
delta2 <- 0.25

set.seed(1234)
z2 <- rnormratio(800, bet = beta2, rho = rho2, delta = delta2)

# With the EM method:
estparnormratio(z2, method = "EM")

# With the variational method:
```

```
estparnormratio(z2, method = "VB")
```

`kummer` *Confluent D-Hypergeometric Function*

Description

Computes the Kummer's function, or confluent hypergeometric function.

Usage

```
kummer(a, b, z, eps = 1e-06)
```

Arguments

`a`	numeric.
`b`	numeric
`z`	numeric vector.
`eps`	numeric. Precision for the sum (default 1e-06).

Details

The Kummer's confluent hypergeometric function is given by:

$$ {}_1F_1\left(a, b; z\right) = \sum_{n=0}^{+\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!} $$

where $(z)_p$ is the Pochhammer symbol (see `pochhammer`).

The `eps` argument gives the required precision for its computation. It is the `attr(, "epsilon")` attribute of the returned value.

Value

A numeric value: the value of the Kummer's function, with two attributes `attr(, "epsilon")` (precision of the result) and `attr(, "k")` (number of iterations).

Author(s)

Pierre Santagostini, Angélina El Ghaziri, Nizar Bouhlel

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

`lnpochhammer` *Logarithm of the Pochhammer Symbol*

Description

Computes the logarithm of the Pochhammer symbol.

Usage

```
lnpochhammer(x, n)
```

Arguments

`x`	numeric.
`n`	positive integer.

Details

The Pochhammer symbol is given by:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)...(x+n-1)$$

So, if $n > 0$:

$$log\left((x)_n\right) = log(x) + log(x+1) + ... + log(x+n-1)$$

If $n = 0$, $log\left((x)_n\right) = log(1) = 0$

Value

Numeric value. The logarithm of the Pochhammer symbol.

Author(s)

Pierre Santagostini, Nizar Bouhlel

See Also

`pochhammer`, `kummer`

Examples

```
lnpochhammer(2, 0)
lnpochhammer(2, 1)
lnpochhammer(2, 3)
```

pnormratio *Cumulative Distribution of a Normal Ratio Distribution*

Description

Cumulative distribution of the ratio of two independent Gaussian distributions.

Usage

```
pnormratio(z, bet, rho, delta)
```

Arguments

`z`	length p vector of quantiles.
`bet, rho, delta`	numeric values. The parameters (β, ρ, δ_y) of the distribution, see Details.

Details

Let two independant random variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$.

If we denote $f_Z(z; \beta, \rho, \delta_y)$ the probability distribution function of the ratio $Z = \frac{X}{Y}$, with $\beta = \frac{\mu_x}{\mu_y}$, $\rho = \frac{\sigma_y}{\sigma_x}$ and $\delta_y = \frac{\sigma_y}{\mu_y}$ (see `dnormratio()`, Details section).

The probability distribution for Z is given by:

$$F(z; \beta, \rho, \delta_y) = \int_{-\infty}^{z} f_Z(z; \beta, \rho, \delta_y)$$

This integral is computed using numerical integration.

Value

Numeric: the value of density.

Author(s)

Pierre Santagostini, Angélina El Ghaziri, Nizar Bouhlel

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

Marsaglia, G. 2006. Ratios of Normal Variables. Journal of Statistical Software 16. doi:10.18637/jss.v016.i04

Díaz-Francés, E., Rubio, F.J., On the existence of a normal approximation to the distribution of the ratio of two independent normal random variables. Stat Papers 54, 309–323 (2013). doi:10.1007/s0036201204292

See Also

`dnormratio()`: density function.

`rnormratio()`: sample simulation.

`estparnormratio()`: parameter estimation.

Examples

```
# First example
beta1 <- 0.15
rho1 <- 5.75
delta1 <- 0.22
pnormratio(0, bet = beta1, rho = rho1, delta = delta1)
pnormratio(0.5, bet = beta1, rho = rho1, delta = delta1)
curve(pnormratio(x, bet = beta1, rho = rho1, delta = delta1), from = -0.1, to = 0.7)

# Second example
beta2 <- 2
rho2 <- 2
delta2 <- 2
pnormratio(0, bet = beta2, rho = rho2, delta = delta2)
pnormratio(0.5, bet = beta2, rho = rho2, delta = delta2)
curve(pnormratio(x, bet = beta2, rho = rho2, delta = delta2), from = -0.1, to = 0.7)
```

---

pochhammer — *Pochhammer Symbol*

Description

Computes the Pochhammer symbol.

Usage

```
pochhammer(x, n)
```

Arguments

x	numeric.
n	positive integer.

Details

The Pochhammer symbol is given by:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)...(x+n-1)$$

Value

Numeric value. The value of the Pochhammer symbol.

Author(s)

Pierre Santagostini, Nizar Bouhlel

See Also

lnpochhammer, kummer

Examples

```
pochhammer(2, 0)
pochhammer(2, 1)
pochhammer(2, 3)
```

---

rnormratio *Simulate from a Normal Ratio Distribution*

Description

Simulate data from a ratio of two independent Gaussian distributions.

Usage

```
rnormratio(n, bet, rho, delta)
```

Arguments

n	integer. Number of observations. If length(n) > 1, the length is taken to be the nmber required.
bet, rho, delta	numeric values. The parameters (β, ρ, δ_y) of the distribution, see Details.

Details

Let two random variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$

with probability densities f_X and f_Y.

The parameters of the distribution of the ratio $Z = \frac{X}{Y}$ are: $\beta = \dfrac{\mu_x}{\mu_y}$, $\rho = \dfrac{\sigma_y}{\sigma_x}$, $\delta_y = \dfrac{\sigma_y}{\mu_y}$.

μ_x, σ_x, μ_y and σ_y are computed from β, ρ and δ_y (by fixing arbitrarily $\mu_x = 1$) and two random samples $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are simulated.

Then $\left(\dfrac{x_1}{y_1}, \ldots, \dfrac{x_n}{y_n}\right)$ is returned.

Value

A numeric vector: the produced sample.

Author(s)

Pierre Santagostini, Angélina El Ghaziri, Nizar Bouhlel

References

El Ghaziri, A., Bouhlel, N., Sapoukhina, N., Rousseau, D., On the importance of non-Gaussianity in chlorophyll fluorescence imaging. Remote Sensing 15(2), 528 (2023). doi:10.3390/rs15020528

Marsaglia, G. 2006. Ratios of Normal Variables. Journal of Statistical Software 16. doi:10.18637/jss.v016.i04

Díaz-Francés, E., Rubio, F.J., On the existence of a normal approximation to the distribution of the ratio of two independent normal random variables. Stat Papers 54, 309–323 (2013). doi:10.1007/s0036201204292

See Also

`dnormratio()`: probability density of a normal ratio.

`pnormratio()`: probability distribution function.

`estparnormratio()`: parameter estimation.

Examples

```
# First example
beta1 <- 0.15
rho1 <- 5.75
delta1 <- 0.22
rnormratio(20, bet = beta1, rho = rho1, delta = delta1)

# Second example
beta2 <- 0.24
rho2 <- 4.21
delta2 <- 0.25
rnormratio(20, bet = beta2, rho = rho2, delta = delta2)
```

Index