Package ‘RMTstat’

October 12, 2022

Version 0.3.1
Date 2022-04-12
Title Distributions, Statistics and Tests Derived from Random Matrix Theory
Author Iain M. Johnstone, Zongming Ma, Patrick O. Perry, Morteza Shahram, and Evan Biederstedt
Maintainer Evan Biederstedt <evan.biederstedt@gmail.com>
Description Functions for working with the Tracy-Widom laws and other distributions related to the eigenvalues of large Wishart matrices. The tables for computing the Tracy-Widom densities and distribution functions were computed by functions were computed by Momar Dieng’s MATLAB package `RMLab`. This package is part of a collaboration between Iain Johnstone, Zongming Ma, Patrick Perry, and Morteza Shahram.
License BSD_3_clause + file LICENSE
Encoding UTF-8
URL https://github.com/evanbiederstedt/RMTstat
BugReports https://github.com/evanbiederstedt/RMTstat
NeedsCompilation no
Repository CRAN
Date/Publication 2022-04-12 23:02:34 UTC

R topics documented:

MarcenkoPastur .......................................................... 2
TracyWidom ............................................................ 3
WishartMax .............................................................. 4
WishartMaxPar .......................................................... 6
WishartSpike ............................................................ 7
WishartSpikePar ........................................................ 8

Index 10
Description

Density, distribution function, quantile function and random generation for the Marčenko-Pastur distribution, the limiting distribution of the empirical spectral measure for a large white Wishart matrix.

Usage

\[dmp( x, ndf=NA, pdim=NA, var=1, svr=ndf/pdim, log = FALSE )\]
\[pmp( q, ndf=NA, pdim=NA, var=1, svr=ndf/pdim, lower.tail = TRUE, log.p = FALSE )\]
\[qmp( p, ndf=NA, pdim=NA, var=1, svr=ndf/pdim, lower.tail = TRUE, log.p = FALSE )\]
\[rmp( n, ndf=NA, pdim=NA, var=1, svr=ndf/pdim )\]

Arguments

- \(x, q\): vector of quantiles.
- \(p\): vector of probabilities.
- \(n\): number of observations. If \(\text{length}(n) > 1\), the length is taken to be the number required.
- \(ndf\): the number of degrees of freedom for the Wishart matrix.
- \(pdim\): the number of dimensions (variables) for the Wishart matrix.
- \(var\): the population variance.
- \(svr\): samples to variables ratio; the number of degrees of freedom per dimension.
- \(log, log.p\): logical; if TRUE, probabilities \(p\) are given as \(\log(p)\).
- \(lower\_tail\): logical; if TRUE (default), probabilities are \(P[X \leq x]\), otherwise, \(P[X > x]\).

Details

The concentration can either be given explicitly, or else computed from the given \(ndf\) and \(pdim\). If \(var\) is not specified, it assumes the default of 1.

The Marčenko-Pastur law is the limit of the random probability measure which puts equal mass on all \(pdim\) eigenvalues of a normalized \(pdim\)-dimensional white Wishart matrix with \(ndf\) degrees of freedom and scale parameter \(\text{diag}(var, var, \ldots, var)\). It is assumed that \(ndf\) goes to infinity, and \(ndf/pdim\) goes to nonzero constant called the "samples-to-variables ratio" (\(svr\)).

Value

- \(dmp\) gives the density, \(pmp\) gives the distribution function, \(qmp\) gives the quantile function, and \(rmp\) generates random deviates.

Author(s)

Iain M. Johnstone, Zongming Ma, Patrick O. Perry and Morteza Shahram
**Source**

Other than the density, these functions are relatively slow and imprecise.

The distribution function is computed with `integrate`. The quantiles are computed via bisection using `uniroot`. Random variates are generated using the inverse CDF.

**References**


---

**The Tracy-Widom Distributions**

**Description**

Density, distribution function, quantile function, and random generation for the Tracy-Widom distribution with order parameter $\beta$.

**Usage**

```
dtw(x, beta=1, log = FALSE)
ptw(q, beta=1, lower.tail = TRUE, log.p = FALSE)
qtw(p, beta=1, lower.tail = TRUE, log.p = FALSE)
rtw(n, beta=1)
```

**Arguments**

- `x,q` vector of quantiles.
- `p` vector of probabilities.
- `n` number of observations. If `length(n) > 1`, the length is taken to be the number required.
- `beta` the order parameter (1, 2, or 4).
- `log, log.p` logical; if TRUE, probabilities p are given as log(p).
- `lower.tail` logical; if TRUE (default), probabilities are $P[X \leq x]$, otherwise, $P[X > x]$.

**Details**

If `beta` is not specified, it assumes the default value of 1.

The Tracy-Widom law is the edge-scaled limiting distribution of the largest eigenvalue of a random matrix from the $\beta$-ensemble. Supported values for `beta` are 1 (Gaussian Orthogonal Ensemble), 2 (Gaussian Unitary Ensemble), and 4 (Gaussian Symplectic Ensemble).

**Value**

- `dtw` gives the density,
- `ptw` gives the distribution function,
- `qtw` gives the quantile function,
- `rtw` generates random deviates.
The distribution and density functions are computed using a lookup table. They have been pre-computed at 769 values uniformly spaced between -10 and 6 using MATLAB’s bvp4c solver to a minimum accuracy of about 3.4e-08. For all other points, the values are gotten from a cubic Hermite polynomial interpolation. The MATLAB software for computing the grid of values is part of RMLab, a package written by Momar Dieng.

The quantiles are computed via bisection using uniroot.

Random variates are generated using the inverse CDF.

References


The White Wishart Maximum Eigenvalue Distributions

Description

Density, distribution function, quantile function, and random generation for the maximum eigenvalue from a white Wishart matrix (sample covariance matrix) with ndf degrees of freedom, pdim dimensions, population variance var, and order parameter beta.

Usage

dWishartMax(x, ndf, pdim, var=1, beta=1, log = FALSE)
pWishartMax(q, ndf, pdim, var=1, beta=1, lower.tail = TRUE, log.p = FALSE)
qWishartMax(p, ndf, pdim, var=1, beta=1, lower.tail = TRUE, log.p = FALSE)
rWishartMax(n, ndf, pdim, var=1, beta=1)

Arguments

x, q         vector of quantiles.

p           vector of probabilities.

n           number of observations. If length(n) > 1, the length is taken to be the number required.

ndf         the number of degrees of freedom for the Wishart matrix
WishartMax

- **pdim**: The number of dimensions (variables) for the Wishart matrix.
- **var**: The population variance.
- **beta**: The order parameter (1 or 2).
- **log, log.p**: Logical; if TRUE, probabilities p are given as log(p).
- **lower.tail**: Logical; if TRUE (default), probabilities are \( P[X \leq x] \), otherwise, \( P[X > x] \).

**Details**

If **beta** is not specified, it assumes the default value of 1. Likewise, **var** assumes a default of 1.

A white Wishart matrix is equal in distribution to \((1/n)X'X\), where \(X\) is an \(n \times p\) matrix with elements i.i.d. Normal with mean zero and variance **var**. These functions give the limiting distribution of the largest eigenvalue from such a matrix when ndf and **pdim** both tend to infinity.

Supported values for **beta** are 1 for real data and 2 for complex data.

**Value**

- `dWishartMax` gives the density, `pWishartMax` gives the distribution function, `qWishartMax` gives the quantile function, and `rWishartMax` generates random deviates.

**Author(s)**

Iain M. Johnstone, Zongming Ma, Patrick O. Perry and Morteza Shahram

**Source**

The functions are calculated by applying the appropriate centering and scaling (determined by `WishartMaxPar`), and then calling the corresponding functions for the `TracyWidom` distribution.

**References**


**See Also**

`WishartMaxPar`, `WishartSpike`, `TracyWidom`
Description

Centering and scaling for the maximum eigenvalue from a white Wishart matrix (sample covariance matrix) with with \( \text{ndf} \) degrees of freedom, \( \text{pdim} \) dimensions, population variance \( \text{var} \), and order parameter \( \text{beta} \).

Usage

\[
\text{WishartMaxPar}(\text{ndf}, \text{pdim}, \text{var}=1, \text{beta}=1)
\]

Arguments

- \( \text{ndf} \): the number of degrees of freedom for the Wishart matrix.
- \( \text{pdim} \): the number of dimensions (variables) for the Wishart matrix.
- \( \text{var} \): the population variance.
- \( \text{beta} \): the order parameter (1 or 2).

Details

If \( \text{beta} \) is not specified, it assumes the default value of 1. Likewise, \( \text{var} \) assumes a default of 1.

The returned values give appropriate centering and scaling for the largest eigenvalue from a white Wishart matrix so that the centered and scaled quantity converges in distribution to a Tracy-Widom random variable. We use the second-order accurate versions of the centering and scaling given in the references below.

Value

- \( \text{centering} \): gives the centering.
- \( \text{scaling} \): gives the scaling.

Author(s)

Iain M. Johnstone, Zongming Ma, Patrick O. Perry and Morteza Shahram

References


See Also

- *WishartMax*, *TracyWidom*
WishartSpike

The Spiked Wishart Maximum Eigenvalue Distributions

Description

Density, distribution function, quantile function, and random generation for the maximum eigenvalue from a spiked Wishart matrix (sample covariance matrix) with ndf degrees of freedom, pdim dimensions, and population covariance matrix diag(spike+var, var, var,..., var).

Usage

dWishartSpike(x, spike, ndf=NA, pdim=NA, var=1, beta=1, log = FALSE)
pWishartSpike(q, spike, ndf=NA, pdim=NA, var=1, beta=1, lower.tail = TRUE, log.p = FALSE)
qWishartSpike(p, spike, ndf=NA, pdim=NA, var=1, beta=1, lower.tail = TRUE, log.p = FALSE)
rWishartSpike(n, spike, ndf=NA, pdim=NA, var=1, beta=1)

Arguments

  x, q        vector of quantiles.
  p           vector of probabilities.
  n           number of observations. If length(n) > 1, the length is taken to be the number required.
  spike       the value of the spike.
  ndf         the number of degrees of freedom for the Wishart matrix.
  pdim        the number of dimensions (variables) for the Wishart matrix.
  var         the population (noise) variance.
  beta        the order parameter (1 or 2).
  log, log.p  logical; if TRUE, probabilities p are given as log(p).
  lower.tail  logical; if TRUE (default), probabilities are P[X ≤ x], otherwise, P[X > x].

Details

The spiked Wishart is a random sample covariance matrix from multivariate normal data with ndf observations in pdim dimensions. The spiked Wishart has one population covariance eigenvalue equal to spike+var and the rest equal to var. These functions are related to the limiting distribution of the largest eigenvalue from such a matrix when ndf and pdim both tending to infinity, with their ratio tending to a nonzero constant.

For the spiked distribution to exist, spike must be greater than sqrt(pdim/ndf)*var.

Supported values for beta are 1 for real data and and 2 for complex data.

Value

dWishartSpike gives the density, pWishartSpike gives the distribution function, qWishartSpike gives the quantile function, and rWishartSpike generates random deviates.
Author(s)
Iain M. Johnstone, Zongming Ma, Patrick O. Perry and Morteza Shahram

References

See Also
WishartSpikePar, WishartMax

---

WishartSpikePar

**Spiked Wishart Eigenvalue Centering and Scaling**

**Description**
Centering and scaling for the sample eigenvalue from a spiked Wishart matrix (sample covariance matrix) with \( \text{ndf} \) degrees of freedom, \( \text{pdim} \) dimensions, and population covariance matrix \( \text{diag}(\text{spike+var}, \text{var}, \text{var}, \ldots, \text{var}) \).

**Usage**
```
WishartSpikePar( spike, ndf=NA, pdim=NA, var=1, beta=1 )
```

**Arguments**
- **spike**: the value of the spike.
- **ndf**: the number of degrees of freedom for the Wishart matrix.
- **pdim**: the number of dimensions (variables) for the Wishart matrix.
- **var**: the population (noise) variance.
- **beta**: the order parameter (1 or 2).

**Details**
The returned values give appropriate centering and scaling for the largest eigenvalue from a spiked Wishart matrix so that the centered and scaled quantity converges in distribution to a normal random variable with mean 0 and variance 1.
For the spiked distribution to exist, \( \text{spike} \) must be greater than \( \sqrt{\text{pdim}/\text{ndf}} \times \text{var} \).
Supported values for \( \beta \) are 1 for real data and 2 for complex data.
WishartSpikePar

Value

centering gives the centering.
scaleing gives the scaling.

Author(s)
Iain M. Johnstone, Zongming Ma, Patrick O. Perry and Morteza Shahram

References

See Also
WishartSpike
Index

* distribution
  MarcenkoPastur, 2
  TracyWidom, 3
  WishartMax, 4
  WishartMaxPar, 6
  WishartSpike, 7
  WishartSpikePar, 8

dmp (MarcenkoPastur), 2
dtw (TracyWidom), 3
dWishartMax (WishartMax), 4
dWishartSpike (WishartSpike), 7

integrate, 3

MarcenkoPastur, 2

pmp (MarcenkoPastur), 2
ptw (TracyWidom), 3
pWishartMax (WishartMax), 4
pWishartSpike (WishartSpike), 7

qmp (MarcenkoPastur), 2
qtw (TracyWidom), 3
qWishartMax (WishartMax), 4
qWishartSpike (WishartSpike), 7

rmp (MarcenkoPastur), 2
rtw (TracyWidom), 3
rWishartMax (WishartMax), 4
rWishartSpike (WishartSpike), 7

TracyWidom, 3, 5, 6

unroot, 3, 4

WishartMax, 4, 6, 8
WishartMaxPar, 5, 6
WishartSpike, 5, 7, 9
WishartSpikePar, 8, 8