

# Package: SOPIE (via r-universe)

September 9, 2024

**Type** Package

**Title** Non-Parametric Estimation of the Off-Pulse Interval of a Pulsar

**Version** 1.6

**Date** 2022-02-23

**Author** Willem Daniel Schutte

**Maintainer** Willem Daniel Schutte <wd.schutte@nwu.ac.za>

**Depends** R (>= 4.0), circular, ADGofTest, grDevices, graphics

**Description** Provides functions to non-parametrically estimate the off-pulse interval of a source function originating from a pulsar. The technique is based on a sequential application of P-values obtained from goodness-of-fit tests for the uniform distribution, such as the Kolmogorov-Smirnov, Cramer-von Mises, Anderson-Darling and Rayleigh goodness-of-fit tests.

**License** GPL-3

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2022-02-23 16:30:06 UTC

## Contents

SOPIE-package . . . . .	2
a.estimate . . . . .	3
b.estimate . . . . .	5
circ.kernel . . . . .	7
crab . . . . .	8
findh . . . . .	9
J1709 . . . . .	10
simdata . . . . .	11
SOPIE . . . . .	12
von_mises_sim . . . . .	14

<b>Index</b>	<b>16</b>
--------------	-----------

## Description

The package 'SOPIE' provides functions to non-parametrically estimate the off-pulse interval of a source function originating from a pulsar. This technique is based on a sequential application of P-values obtained from goodness-of-fit tests for the uniform distribution. The well-known Kolmogorov-Smirnov, Cramer-von Mises, Anderson-Darling and Rayleigh test statistics are applied sequentially on subintervals of  $[0; 1]$ .

The most important functions in the package are combined in a wrapper function called [SOPIE](#). Users should start by looking at the documentation of the function [findh](#), [circ.kernel](#) and [SOPIE](#).

## Details

Package:	SOPIE
Type:	Package
Version:	1.6
Date:	2022-02-23
License:	GPL-3
LazyLoad:	yes

The SOPIE package consists of 4 main functions. Each of these functions are discussed in terms of its functioning, structure, arguments and output in the help documentation of each function.

1. [findh](#) is the function used to obtain the estimated smoothing parameter  $\hat{h}$  that will be used in the circular kernel density estimator.
2. [circ.kernel](#) is the function used to perform circular kernel density estimation on the sample data set in order to obtain the minimum points of the kernel density estimator. This is essentially the first step of the suggested procedure, as described in the second reference listed below. The output can also be used to draw a graph of the circular kernel density estimator.
3. [a.estimate](#) and [b.estimate](#) is almost identical function. [a.estimate](#) is the function used to obtain the estimated values of  $a$ , i.e.  $\hat{a}$ , for the off-pulse interval of a pulsar light curve. [b.estimate](#) is the function used to obtain the estimated values of  $b$ , i.e.  $\hat{b}$ , for the off-pulse interval of a pulsar light curve.
4. [SOPIE](#) is a wrapper-function in the sense that it utilises all of the above function to produce the estimated off-pulse intervals in an easy readable matrix format, together with a graph consisting of the histogram estimate of the sample data, the kernel density estimator and an indication of the estimated median off-pulse interval.

## Author(s)

Willem Daniel Schutte  
 Maintainer: Willem Daniel Schutte

## References

- Jammalamadaka, S. Rao and SenGupta, A. (2001). Topics in Circular Statistics, World Scientific Publishing Co. Pte. Ltd.
- Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. Monthly Notices of the Royal Astronomical Society, 461, 627-640.

## Examples

```
set.seed(777)
simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
SOPIE(simdata,h=1,to=1,alpha=0.05,g=5,r=10,m=1,grid=100)
```

---

a.estimate

*Estimate the Left Endpoint of the Off-Pulse Interval of a Pulsar*

---

## Description

a.estimate and b.estimate is almost identical function. a.estimate is the function used to obtain the estimated values of  $a$ , i.e.  $\hat{a}$ , for the off-pulse interval of a pulsar light curve. b.estimate is the function used to obtain the estimated values of  $b$ , i.e.  $\hat{b}$ , for the off-pulse interval of a pulsar light curve.

## Usage

```
a.estimate(data, to = 1, min_points, alpha = 0.05, g = 1, r = 1)
```

## Arguments

data	the data vector used to estimate $a$ .
to	the value of the maximum domain of the data. Values will usually either be 1 or $2\pi$ .
min_points	the scalar or vector containing the value(s) of the minimum point(s) calculated during the kernel density estimation. This argument does not represent the index value(s) of the observations within data. The minimum point(s) can be obtained with the function <a href="#">circ.kernel</a> .
alpha	significance level ( $\alpha$ ) that will be used during the sequential application of the goodness-of-fit tests for uniformity when estimating the off-pulse interval.
g	the value of the incremental growth of each subsequent interval over which uniformity is tested. In the suggested procedure, uniformity is sequentially tested, with the interval used in the test growing by $g$ observations after every iteration. The selection of $g$ not only influences the computation time of the procedure, but

also has an effect on the point where rejection of the hypothesis takes place. For large values of  $g$ , the user takes the risk that uniformity is rejected for a certain (larger) interval, while it should have been rejected earlier (for a smaller interval). On the other hand, a very small choice of  $g$  results in long execution times. Small values of  $g$  may also result in the early rejection of uniformity, e.g. in the situation where a few observations may cause the rejection of uniformity, while uniformity is again confirmed when several more observations are included in the interval. If the user suspects that this situation may occur, the problem can be overcome by selecting a larger value of the integer  $r$ .

$r$  the number of subsequent intervals that must result in the rejection of uniformity before the function will stop. The choice of  $r$  must therefore be linked to the choice of  $g$  as explained above. For smaller values of  $g$ , it would be safer to select larger values of  $r$ , and vice versa. Since small values of  $g$  may result in a temporary rejection of uniformity for an interval, a larger value of  $r$  would prevent the method from immediately stopping at the first occurrence of rejection. It is very important to note that, for a large value of  $r$ , there will be no impact on the value of  $\hat{b}$  or  $\hat{a}$  if rejection takes place for each interval after a certain point.

### Value

a list containing the following components:

summary	a vector containing the estimated value of $a$ , i.e. $\hat{a}$ , for each of the four goodness-of-fit tests, namely the Anderson-Darling, Kolmogorov-Smirnov, Cramer-von Mises and the Rayleigh goodness-of-fit test.
general	a list containing the function call, the minimum value(s) used in the estimation, the level of significance ( $\alpha$ ), the value of $g$ and the value of $r$ .

### Author(s)

Willem Daniel Schutte

### References

- D'Agostino, R. & Stephens, M. (eds) (1986). Goodness-of-fit techniques, Marcel Dekker, Inc.
- Jammalamadaka, S. Rao and SenGupta, A. (2001). Topics in Circular Statistics, World Scientific Publishing Co. Pte. Ltd.
- Marsaglia G, Marsaglia J (2004). Evaluating the Anderson-Darling Distribution. Journal of Statistical software, 9, 1-5.
- Marsaglia G, Tsang WW, Wang J (2003). Evaluating Kolmogorov's Distribution. Journal of Statistical Software, 8(18), 1-4.
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. Monthly Notices of the Royal Astronomical Society, 461, 627-640.
- Stephens M (1970). Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables. Journal of the Royal Statistical Society. Series B (Methodological), 32, 115-122.

**See Also**

[ad.test](#) , [ks.test](#) , [rayleigh.test](#)

**Examples**

```
## This function is to be used inside the wrapper function SOPIE

simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
SOPIE(simdata,h=1,to=1,alpha=0.05,g=5,r=10,m=1,grid=100)
```

---

b.estimate

*Estimate the Right Endpoint of the Off-Pulse Interval of a Pulsar*


---

**Description**

a.estimate and b.estimate is almost identical function. a.estimate is the function used to obtain the estimated values of  $a$ , i.e.  $\hat{a}$  for the off-pulse interval of a pulsar light curve. b.estimate is the function used to obtain the estimated values of  $b$ , i.e.  $\hat{b}$ , for the off-pulse interval of a pulsar light curve.

**Usage**

```
b.estimate(data, to = 1, min_points, alpha = 0.05, g = 1, r = 1)
```

**Arguments**

data	the data vector used to estimate $b$ .
to	the value of the maximum domain of the data. Values will usually either be 1 or $2\pi$ .
min_points	a scalar or vector containing the value(s) of the minimum point(s) calculated during the kernel density estimation. This argument does not represent the index value(s) of the observations within data. The minimum point(s) can be obtained with the function <a href="#">circ.kernel</a> .
alpha	significance level ( $\alpha$ ) that will be used during the sequential application of the goodness-of-fit tests for uniformity when estimating the off-pulse interval.
g	the value of the incremental growth of each subsequent interval over which uniformity is tested. In the suggested procedure, uniformity is sequentially tested, with the interval used in the test growing by $g$ observations after every iteration. The selection of $g$ not only influences the computation time of the procedure, but also has an effect on the point where rejection of the hypothesis takes place. For large values of $g$ , the user takes the risk that uniformity is rejected for a certain (larger) interval, while it should have been rejected earlier (for a smaller interval). On the other hand, a very small choice of $g$ results in long execution times. Small values of $g$ may also result in the early rejection of uniformity, e.g. in the situation where a few observations may cause the rejection of uniformity, while uniformity is again confirmed when several more observations are included in

the interval. If the user suspects that this situation may occur, the problem can be overcome by selecting a larger value of the integer  $r$ .

$r$  the number of subsequent intervals that must result in the rejection of uniformity before the function will stop. The choice of  $r$  must therefore be linked to the choice of  $g$  as explained above. For smaller values of  $g$ , it would be safer to select larger values of  $r$ , and vice versa. Since small values of  $g$  may result in a temporary rejection of uniformity for an interval, a larger value of  $r$  would prevent the method from immediately stopping at the first occurrence of rejection. It is very important to note that, for a large value of  $r$ , there will be no impact on the value of  $\hat{b}$  or  $\hat{a}$  if rejection takes place for each interval after a certain point.

### Value

a list containing the following components:

summary	a vector containing the estimated value of $b$ , i.e. $\hat{b}$ , for each of the four goodness-of-fit tests, namely the Anderson-Darling, Kolmogorov-Smirnov, Cramer-von Mises and the Rayleigh goodness-of-fit test.
\$general	a list containing the function call, the minimum value(s) used in the estimation, the level of significance ( $\alpha$ ), the value of $g$ and the value of $r$ .

### Author(s)

Willem Daniel Schutte

### References

- D'Agostino, R. & Stephens, M. (eds) (1986). Goodness-of-fit techniques, Marcel Dekker, Inc.
- Jammalamadaka, S. Rao and SenGupta, A. (2001). Topics in Circular Statistics, World Scientific Publishing Co. Pte. Ltd.
- Marsaglia G, Marsaglia J (2004). Evaluating the Anderson-Darling Distribution. Journal of Statistical software, 9, 1-5.
- Marsaglia G, Tsang WW, Wang J (2003). Evaluating Kolmogorov's Distribution. Journal of Statistical Software, 8(18), 1-4.
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. Monthly Notices of the Royal Astronomical Society, 461, 627-640.
- Stephens M (1970). Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables. Journal of the Royal Statistical Society. Series B (Methodological), 32, 115-122.

### See Also

[ad.test](#) , [ks.test](#) , [rayleigh.test](#)

**Examples**

```
## This function is to be used inside the wrapper function SOPIE

simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
SOPIE(simdata,h=1,to=1,alpha=0.05,g=5,r=10,m=1,grid=100)
```

---

circ.kernel

*Circular Kernel Density Estimation*


---

**Description**

This function is used to perform circular kernel density estimation on the sample data set in order to obtain the minimum points of the kernel density estimator.

**Usage**

```
circ.kernel(data, sp, to = 1, grid = 512, m = 1)
```

**Arguments**

data	the data vector from which the circular kernel density estimator is to be computed.
sp	a real value ( $0 < sp < 1$ ) for the smoothing parameter to be used. This value can be obtained by using <a href="#">findh</a> .
to	the value of the maximum domain of the data. Values will usually either be 1 or $2\pi$ .
grid	the number of equally spaced grid points at which the density is to be estimated.
m	the number of local minimum points included in the output.

**Details**

The Epanechnikov kernel function is used in the circular kernel density estimation. Circular kernel density estimation is performed according to the method proposed in 'Topics in circular statistics' (see references).

**Value**

a list containing the following components:

x	a vector of sorted $x$ values that represents the equally-spaced grid points used during the kernel density estimation.
y	a vector of density-values of the circular kernel density estimator corresponding to $x$ .
minimum	a vector of the kernel grid point(s) of lowest density derived from the circular kernel density estimator. The length of the vector will depend on the choice of $m$ .

**Author(s)**

Willem Daniel Schutte

**References**

- Hall P, Watson G, Cabrera J (1987). Kernel density estimation with spherical data. *Biometrika*, 74 (4), 751-762.
- Jammalamadaka S, SenGupta A (2001). *Topics in circular statistics*. World Scientific Publishing Co. Pte. Ltd.
- Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. *Monthly Notices of the Royal Astronomical Society*, 461, 627-640.
- Sheather, S. & Jones, M. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B*, 53:683-690.
- Silverman, B. (1986). *Density estimation for Statistics and Data analysis*, Chapman and Hall.
- Taylor, C. (2008). Automatic bandwidth selection for circular density estimation, *Computational Statistics & Data Analysis*, 52:3493-3500.
- Wand, M. & Jones, M. (1995). *Kernel Smoothing*, Chapman and Hall.

**Examples**

```
simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
circ.kernel(simdata, findh(simdata), to = 1, grid = 512, m = 1)
```

---

crab

*PSR J0534+2200 (Crab-Pulsar) Time of Arrivals*

---

**Description**

This data set contains n=21145 time of arrivals of photons with energies above 100MeV of PSR J0534+2200 (Crab-pulsar), obtained from the Fermi LAT.

**Usage**

```
data(crab)
```

**Format**

A vector containing 21145 observation.

**Source**

Obtained from Fermi LAT, energies above 100 MeV.



## References

Abdo A, et al. (2010b). Fermi large area telescope observations of the Crab pulsar and nebula. The Astronomical Journal, 708, 1254-1267.

## Examples

```
data(crab)
SOPIE(crab)
```

---

findh	<i>Calculate the Estimated Smoothing Parameter</i>
-------	--

---

## Description

This function is used to obtain the estimated smoothing parameter  $\hat{h}$  that will be used in the circular kernel density estimator (see [circ.kernel](#)).

## Usage

```
findh(data, h = 1, to = 1)
```

## Arguments

data	the data vector from which to calculate the estimated smoothing parameter $\hat{h}$ that will be used in the circular kernel density estimator.
h	integer value from 1 to 9, specifying the smoothing parameter to calculate according to the following table:

$$\begin{aligned}\hat{h}_1 &= 1.06sn^{-1/5} \\ \hat{h}_2 &= 1.06s_0n^{-1/5} \\ \hat{h}_3 &= 1.06\bar{D}_0n^{-1/5} \\ \hat{h}_4 &= 1.06|D_0|n^{-1/5} \\ \hat{h}_5 &= 1.06IQR_0n^{-1/5} \\ \hat{h}_6 &= \frac{1.06}{1.349}IQR_0n^{-1/5} \\ \hat{h}_7 &= 0.9s_0n^{-1/5} \\ \hat{h}_8 &= \frac{0.9}{1.349}IQR_0n^{-1/5} \\ \hat{h}_9 &= \frac{1}{8}\sum_{i=1}^8 h_i\end{aligned}$$

to	the value of the maximum domain of the data. Values will usually either be 1 or $2\pi$ .
----	--

## Value

The function produces a single real value between 0 and 1, representing the rounded value (to 2 decimal places) of the estimating smoothing parameter.

**Author(s)**

Willem Daniel Schutte

**References**

- Hall P, Watson G, Cabrera J (1987). Kernel density estimation with spherical data. *Biometrika*, 74 (4), 751-762.
- Jammalamadaka S, SenGupta A (2001). *Topics in circular statistics*. World Scientific Publishing Co. Pte. Ltd.
- Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. *Monthly Notices of the Royal Astronomical Society*, 461, 627-640.
- Sheather, S. & Jones, M. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B*, 53:683-690.
- Silverman, B. (1986). *Density estimation for Statistics and Data analysis*, Chapman and Hall.
- Taylor, C. (2008). Automatic bandwidth selection for circular density estimation, *Computational Statistics & Data Analysis*, 52:3493-3500.
- Wand, M. & Jones, M. (1995). *Kernel Smoothing*, Chapman and Hall.

**Examples**

```
simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
findh(simdata,h=9,to=1)
```

---

J1709

*PSR J1709-44290 Time of Arrivals*

---

**Description**

This data set contains n=21153 time of arrivals of photons with energies above 100MeV of PSR J1709-44290, obtained from the Fermi LAT.

**Usage**

```
data(J1709)
```

**Format**

A vector containing 21153 observation.

**Source**

Obtained from Fermi LAT, energies above 100 MeV

## References

Abdo A, et al. (2010). "The first Fermi large area telescope catalog of gamma-ray pulsars." The Astrophysical Journal Supplement Series, 187, 460-494.

## Examples

```
data(J1709)
SOPIE(J1709)
```

---

simdata

*Simulated Data from a Scaled Von Mises Distribution with Noise*

---

## Description

This simulated data set contains  $n=5000$  observations from a scaled Von Mises distribution with noise ( $\kappa = 1$ ;  $c = 0.3$ ;  $noise = 0.2$ ). Similar data sets can be generated with the function [von\\_mises\\_sim](#).

## Usage

```
data(simdata)
```

## Format

A vector containing 5000 observations.

## Source

Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>

## Examples

```
data(simdata)
hist(simdata)
SOPIE(simdata)
```

**Description**

SOPIE is a wrapper-function that utilises `findh`, `circ.kernel`, `a.estimate` and `b.estimate` to produce the estimated off-pulse intervals in an easy readable matrix format, together with a graph.

**Usage**

```
SOPIE(data, h = 1, to = 1, alpha = 0.05, g = 20, r = 10, m = 1, grid = 512)
```

**Arguments**

`data` the data vector within which to find the estimated smoothing parameter  $\hat{h}$  that will be used in the circular kernel density estimator. After obtaining the minimum point(s) from the circular kernel density estimator, the estimate off-pulse interval  $[\hat{a}; \hat{b}]$  is given as result.

`h` integer value from 1 to 9, specifying the smoothing parameter to calculate according to the following table:

$$\begin{aligned}\hat{h}_1 &= 1.06sn^{-1/5} \\ \hat{h}_2 &= 1.06s_0n^{-1/5} \\ \hat{h}_3 &= 1.06D_0n^{-1/5} \\ \hat{h}_4 &= 1.06|D_0|n^{-1/5} \\ \hat{h}_5 &= 1.06IQR_0n^{-1/5} \\ \hat{h}_6 &= \frac{1.06}{1.349}IQR_0n^{-1/5} \\ \hat{h}_7 &= 0.9s_0n^{-1/5} \\ \hat{h}_8 &= \frac{0.9}{1.349}IQR_0n^{-1/5} \\ \hat{h}_9 &= \frac{1}{8} \sum_{i=1}^8 h_i\end{aligned}$$

`to` the value of the maximum domain of the data. Values will usually either be 1 or  $2\pi$ .

`alpha` significance level ( $\alpha$ ) that will be used during the sequential application of the goodness-of-fit tests for uniformity when estimating the off-pulse interval.

`g` the value of the incremental growth of each subsequent interval over which uniformity is tested. In the suggested procedure, uniformity is sequentially tested, with the interval used in the test growing by `g` observations after every iteration. The selection of `g` not only influences the computation time of the procedure, but also has an effect on the point where rejection of the hypothesis takes place. For large values of `g`, the user takes the risk that uniformity is rejected for a certain (larger) interval, while it should have been rejected earlier (for a smaller interval). On the other hand, a very small choice of `g` results in long execution times. Small values of `g` may also result in the early rejection of uniformity, e.g. in the situation where a few observations may cause the rejection of uniformity, while

uniformity is again confirmed when several more observations are included in the interval. If the user suspects that this situation may occur, the problem can be overcome by selecting a larger value of the integer  $r$ .

$r$	the number of subsequent intervals that must result in the rejection of uniformity before the function will stop. The choice of $r$ must therefore be linked to the choice of $g$ as explained above. For smaller values of $g$ , it would be safer to select larger values of $r$ , and vice versa. Since small values of $g$ may result in a temporary rejection of uniformity for an interval, a larger value of $r$ would prevent the method from immediately stopping at the first occurrence of rejection. It is very important to note that, for a large value of $r$ , there will be no impact on the value of $\hat{b}$ or $\hat{a}$ if rejection takes place for each interval after a certain point.
$m$	the number of local minimum points included in the output.
grid	the number of equally spaced grid points at which the density is to be estimated.

### Details

SOPIE is a wrapper-function in the sense that it utilises the function `findh`, `circ.kernel`, `a.estimate` and `b.estimate` to produce the estimated off-pulse intervals of a pulsar light curve in an easy readable matrix format, together with a graph consisting of the histogram estimate of the sample data, the kernel density estimator, and a visual representation of the estimated off-pulse intervals.

### Value

The output produced by the function is a list containing the following:

summary	is a matrix that contains the estimated value of $a$ and $b$ for each of the four goodness-of-fit tests, namely the Anderson-Darling, Kolmogorov-Smirnov, Cramer-von Mises and the Rayleigh goodness-of-fit test. Based on the four estimated values of $a$ and $b$ , the median values of $a$ and $b$ are also calculated. This median off-pulse interval is the recommended interval and also the interval that is depicted on the graph.
general	is a list containing the function call, the minimum value(s) used in the estimation, the level of significance ( $\alpha$ ), the value of $g$ and the value of $r$ .

A histogram estimate of the data is produced with the circular kernel density estimate overlaid. An indication of the estimated median off-pulse interval derived from the four goodness-of-fit tests is illustrated with two solid vertical lines.

### Author(s)

Willem Daniel Schutte

### References

- Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>.
- Schutte WD, Swanepoel JWH (2016). SOPIE: an R package for the non-parametric estimation of the off-pulse interval of a pulsar light curve. Monthly Notices of the Royal Astronomical Society, 461, 627-640.

**Examples**

```

set.seed(777)
simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
SOPIE(simdata,h=1,to=1,alpha=0.05,g=5,r=10,m=1,grid=100)
data(crab)
SOPIE(crab)
data(J1709)
SOPIE(J1709)

```

---

von\_mises\_sim

*Generates Simulated Data from a Von Mises Distribution with Noise*


---

**Description**

Generates simulated data over the interval  $[0; 1]$  from a scaled Von Mises distribution with noise.

**Usage**

```
von_mises_sim(n = 5000, k = 1, c = 0.3, noise = 0.2)
```

**Arguments**

n	number of random variates in the simulated data set.
k	concentration parameter $\kappa$ of the Von Mises distribution.
c	the point of truncation of the Von Mises distribution. The value of c represent that value in the interval $[0; c]$ and $[1 - c; 1]$ where the Von Mises density is remove, i.e. $f(\theta) = 0$ for $\theta \in [0; c]$ and $\theta \in [1 - c; 1]$ where $f(\theta)$ is the Von Mises density function.
noise	proportion of random noise to include in the simulated data set. If n random variates are required, then $\lfloor (1 - noise)n \rfloor$ values are generated from the Von Mises density and the remainder from an uniform density.

**Value**

The output vector of this function is  $n$  random variates in the interval  $[0; 1]$  from a scaled Von Mises density with uniform noise proportional to noise.

**Author(s)**

Willem Daniel Schutte

**References**

Jammalamadaka, S. Rao and SenGupta, A. (2001). Topics in Circular Statistics, World Scientific Publishing Co. Pte. Ltd.  
 Robert CP, Casella G (2010). Introducing Monte Carlo methods with R. Springer.  
 Schutte WD (2014). Nonparametric estimation of the off-pulse interval(s) of a pulsar light curve. Ph.D. thesis, North-West University. URL <http://hdl.handle.net/10394/12199>

**See Also**

[pvonmises](#)

**Examples**

```
set.seed(777)
simdata<-von_mises_sim(n=5000,k=1,c=0.3,noise=0.2)
hist(simdata)
SOPIE(simdata,h=1,to=1,alpha=0.05,g=5,r=10,m=1,grid=100)
```

# Index

- \* **datasets**
  - crab, 8
  - J1709, 10
  - simdata, 11
- \* **nonparametric**
  - a.estimate, 3
  - b.estimate, 5
  - SOPIE, 12
  - von\_mises\_sim, 14
- \* **package**
  - SOPIE-package, 2
- \* **smooth**
  - circ.kernel, 7

a.estimate, 2, 3, 12, 13  
ad.test, 5, 6

b.estimate, 2, 5, 12, 13

circ.kernel, 2, 3, 5, 7, 9, 12, 13  
crab, 8

findh, 2, 7, 9, 12, 13

J1709, 10

ks.test, 5, 6

pvonmises, 15

rayleigh.test, 5, 6

simdata, 11  
SOPIE, 2, 12  
SOPIE-package, 2

von\_mises\_sim, 11, 14