

# Package ‘pps’

October 14, 2022

**Version** 1.0

**Date** 2021-01-14

**Title** PPS Sampling

**Author** Jack G. Gambino <jack.gambino@gmail.com>

**Maintainer** Jack G. Gambino <jack.gambino@gmail.com>

**Description** Functions to select samples using PPS (probability proportional to size) sampling. The package also includes a function for stratified simple random sampling, a function to compute joint inclusion probabilities for Sampford's method of PPS sampling, and a few utility functions. The user's guide pps-ug.pdf is included in the ../pps/doc directory. The methods are described in standard survey sampling theory books such as Cochran's "Sampling Techniques"; see the user's guide for references.

**License** GPL (>= 2)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-01-17 05:50:02 UTC

## R topics documented:

calif . . . . .	2
califcty . . . . .	2
permuteinstrata . . . . .	2
pps1 . . . . .	3
ppss . . . . .	3
ppssstrat . . . . .	4
ppswr . . . . .	5
sampford . . . . .	5
sampfordpi . . . . .	6
sizesok . . . . .	7
stratsrs . . . . .	7
stratumsizes . . . . .	8

<b>Index</b>	<b>9</b>
--------------	----------

---

calif	<i>California places</i>
-------	--------------------------

---

**Description**

See user's guide

**Note**

See the user's guide, pps-ug.pdf, for more information.

---

califcty	<i>California counties</i>
----------	----------------------------

---

**Description**

See user's guide

**Note**

See the user's guide, pps-ug.pdf, for more information.

---

permuteinstrata	<i>Randomize units within strata</i>
-----------------	--------------------------------------

---

**Description**

Randomize the order of units within each stratum

**Usage**

```
permuteinstrata(stratsizes)
```

**Arguments**

`stratsizes`      A vector containing the size of each stratum

**Value**

Returns the vector of permuted indices. In the example below, the returned vector has 29 elements.

**Note**

See the user's guide, pps-ug.pdf, for more information.

**Examples**

```
stratsizes <- c(9,10,10) # strata have 9, 10 and 10 units, respectively
permuteinstrata(stratsizes)
```

---

pps1                      *Select one unit with PPS*

---

**Description**

Use PPS systematic sampling to select a single unit out of N

**Usage**

```
pps1(sizes)
```

**Arguments**

sizes                      A vector of the sizes of the units in the population

**Value**

Returns the index of the unit that was selected

**Note**

See the user's guide, pps-ug.pdf, for more information.

**Examples**

```
sizes <- c(9,2,5,17,4,21,15,7,4,11,23,23,14)
sampleindex <- pps1(sizes)
```

---

ppss                      *PPS systematic sampling*

---

**Description**

Use PPS systematic sampling to select a sample of n units out of N

**Usage**

```
ppss(sizes,n)
```

**Arguments**

sizes                      A vector of the sizes of the units in the population  
n                              The sample size

**Value**

Returns the indices of the units that were selected in the sample

**Note**

See the user's guide, pps-ug.pdf, for more information.

**Examples**

```
sizes <- c(9,2,5,17,4,21,15,7,4,11,23,23,14)
sampleindices <- ppss(sizes,4)
```

---

ppssstrat

*Stratified PPS systematic sampling*

---

**Description**

In each stratum, select a sample using pps systematic sampling

**Usage**

```
ppssstrat(sizes, stratum, n)
```

**Arguments**

sizes	A vector of the sizes of the units in the population, sorted by stratum
stratum	A vector of stratum codes, in the same order
n	A vector containing the sample size in each stratum

**Value**

Returns the indices of the units that were selected in the sample

**Note**

ppssstrat calls ppss once per stratum. See the user's guide, pps-ug.pdf, for more information.

**Examples**

```
sizes <- c(1:5,10:6)*10
strat <- c(1,1,1,2,2,3,3,3,3,3)
n <- c(2,1,3)
ppssstrat(sizes, strat, n)
```

---

ppswr

*PPS sampling with replacement*

---

**Description**

Use PPS sampling to select a sample of  $n$  units out of  $N$  with replacement

**Usage**

```
ppswr(sizes, n)
```

**Arguments**

`sizes`            A vector of the sizes of the units in the population  
`n`                 The sample size

**Value**

Returns the indices of the units that were selected in the sample

**Note**

See the user's guide, `pps-ug.pdf`, for more information.

**Examples**

```
sizes <- c(9,2,5,17,4,21,15,7,4,11,23,23,14)  
sampleindices <- ppswr(sizes,4)
```

---

sampford

*Sampford's PPS sampling method*

---

**Description**

Use Sampford's method to select a PPS sample of units

**Usage**

```
sampford(size, n)
```

**Arguments**

`size`            A vector of the sizes of the units in the population  
`n`                 The sample size

**Value**

Returns the indices of the units that were selected in the sample

**Note**

The function `sampfordpi` can be used to compute joint inclusion probabilities for this method. See the user's guide, `pps-ug.pdf`, for more information.

**Examples**

```
size <- c(9,2,5,17,4,21,15,7,4,11,23,23,14)
sampleindices <- sampford(size,4)
```

---

`sampfordpi`

*Joint inclusion probabilities for Sampford's PPS sampling method*

---

**Description**

Compute joint inclusion probabilities for Sampford's method of PPS sampling

**Usage**

```
sampfordpi(sizes,n)
```

**Arguments**

<code>sizes</code>	A vector of the sizes of the units in the population
<code>n</code>	The sample size

**Value**

Returns a matrix with the inclusion probability  $\pi_i(i)$  for each unit  $i$  in the population and with the joint inclusion probability  $\pi_i(i,j)$  of units  $i$  and  $j$  in position  $(i,j)$  in the matrix, where  $i$  and  $j$  are not equal. Note that the size of the matrix is  $N \times N$ , where  $N$  is the population size.

**Note**

The function `sampford` can be used to select a sample using Sampford's method. See the user's guide, `pps-ug.pdf`, for more information.

**Examples**

```
sizes <- c(9,2,5,17,4,21,15,7,4,11,23,23,14)
piij <- sampfordpi(sizes,4)
weights <- 1/diag(piij) # the weights one would use for estimation
```

---

`sizesok`*Check that unit sizes are not too big*

---

**Description**

See user's guide

**Usage**

```
sizesok(size,n)
```

**Arguments**

<code>size</code>	A vector of the sizes of the units in the population
<code>n</code>	The sample size

**Value**

Returns the number of 'bad' units

**Note**

See the user's guide, [pps-ug.pdf](#), for more information.

---

`stratsrs`*Stratified simple random sampling*

---

**Description**

In each stratum, select a simple random sample

**Usage**

```
stratsrs(stratum,nh)
```

**Arguments**

<code>stratum</code>	A vector of stratum codes, sorted by stratum
<code>nh</code>	A vector containing the sample size in each stratum

**Value**

Returns the indices of the units that were selected in the sample

**Note**

See the user's guide, pps-ug.pdf, for more information.

**Examples**

```
strat <- c(1,1,1,1,1,2,2,2,3,3,3,3,3,3,3) # stratum 1 has 5 units, etc.  
nh <- c(2,1,3) # select 2 units from stratum 1, 1 from stratum 2 and 3 from 3  
stratsrs(strat,nh)
```

---

stratumsizes	<i>Compute size of each stratum</i>
--------------	-------------------------------------

---

**Description**

Given a vector of sorted stratum indicators, returns the number of units in each stratum

**Usage**

```
stratumsizes(stratum)
```

**Arguments**

stratum            A vector of sorted stratum indicators

**Value**

Returns the number of units in each stratum

**Note**

See the user's guide, pps-ug.pdf, for more information.



# Index

## \* survey

- calif, 2
- califcty, 2
- permuteinstrata, 2
- pps1, 3
- ppss, 3
- ppssstrat, 4
- ppswr, 5
- sampford, 5
- sampfordpi, 6
- sizesok, 7
- stratsrs, 7
- stratumsizes, 8

- calif, 2
- califcty, 2

- permuteinstrata, 2
- pps1, 3
- ppss, 3
- ppssstrat, 4
- ppswr, 5

- sampford, 5
- sampfordpi, 6
- sizesok, 7
- stratsrs, 7
- stratumsizes, 8