

# Package ‘prnsamplr’

February 9, 2020

**Type** Package

**Title** Permanent Random Number Sampling

**Version** 0.1.1

**Date** 2020-02-09

**Author** Kira Coder Gylling

**Maintainer** Kira Coder Gylling <kira.gylling@gmail.com>

**Description** Survey sampling using permanent random numbers (PRN's). A solution to the problem of unknown overlap between survey samples, which leads to a low precision in estimates when the survey is repeated or combined with other surveys. The PRN solution is to supply the  $U(0, 1)$  random numbers to the sampling procedure, instead of having the sampling procedure generate them. In Lindblom (2014) <doi:10.2478/jos-2014-0047>, and therein cited articles, it is shown how this is carried out and how it improves the estimates. This package supports two common fixed-size sampling procedures (simple random sampling and probability-proportional-to-size sampling) and includes a function for transforming the PRN's in order to control the sample overlap.

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 6.1.1

**Imports** stats

**LazyData** true

**NeedsCompilation** no

**Depends** R (>= 2.10)

**Repository** CRAN

**Date/Publication** 2020-02-09 17:00:02 UTC

## R topics documented:

|                       |   |
|-----------------------|---|
| ExampleData . . . . . | 2 |
| pps . . . . .         | 2 |
| samp . . . . .        | 3 |

|              |                        |            |
|--------------|------------------------|------------|
| 2            |                        | <i>pps</i> |
|              | srs . . . . .          | 4          |
|              | transformprn . . . . . | 5          |
| <b>Index</b> |                        | <b>6</b>   |

|             |                    |
|-------------|--------------------|
| ExampleData | <i>ExampleData</i> |
|-------------|--------------------|

**Description**

Artificial dataset to be used with samp.

**Usage**

```
data("ExampleData")
```

**Format**

A data frame with 40000 observations on the following 6 variables.

- stratum a character vector
- id a numeric vector
- npopul a numeric vector
- nsample a numeric vector
- rands a numeric vector
- sizeM a numeric vector

**Examples**

```
data(ExampleData)
```

|            |   |
|------------|---|
| <i>pps</i> | <i>Stratified probability-proportional-to-size sampling</i> |
|------------|---|

**Description**

Stratified probability-proportional-to-size (PPS) sampling using permanent random numbers. Can also be used for non-stratified PPS using a dummy stratum taking the same value for each object.

**Usage**

```
pps(df, stratid, nsamp, prn, size)
```

**Arguments**

|         |   |
|---------|---|
| df      | Data frame containing the elements to sample from.      |
| stratid | Variable in df containing the strata.                   |
| nsamp   | Variable in df containing the sample sizes.             |
| prn     | Variable in df containing the permanent random numbers. |
| size    | Variable in df containing the size measure.             |

**Value**

Copy of the input data frame together with the boolean variable `sampled`, as well as a numeric variable `lambda` containing the estimated first-order inclusion probabilities.

**See Also**

[samp](#), [srs](#), [transformprn](#).

**Examples**

```
dfOut <- pps(df=ExampleData,
             nsamp="nsample",
             stratid="stratum",
             prn="rands",
             size="sizeM")
```

---

|      |  |
|------|--|
| samp | <i>Stratified permanent random number sampling</i> |
|------|--|

---

**Description**

Wrapper for stratified simple random sampling (SRS) and probability-proportional-to-size (PPS) sampling using permanent random numbers. Can also be used for non-stratified sampling using a dummy stratum taking the same value for each object.

**Usage**

```
samp(method, df, ...)
```

**Arguments**

|        |  |
|--------|--|
| method | pps or srs.  |
| df     | Data frame containing the elements to sample from. |
| ...    | Further method-specific arguments.                 |

**Value**

Copy of the input data frame together with the boolean variable `sampled`, as well as a numeric variable `lambda` containing the estimated first-order inclusion probabilities when PPS is used.

**See Also**

[srs](#), [pps](#), [transformprn](#).

**Examples**

```
dfOut <- samp(method=pps,
              df=ExampleData,
              nsamp="nsample",
              stratid="stratum",
              prn="rands",
              size="sizeM")

dfOut <- samp(method=srs,
              df=ExampleData,
              nsamp="nsample",
              stratid="stratum",
              prn="rands")
```

---

srs

*Stratified simple random sampling*

---

**Description**

Stratified simple random sampling (SRS) using permanent random numbers. Can also be used for non-stratified SRS using a dummy stratum taking the same value for each object.

**Usage**

```
srs(df, stratid, nsamp, prn)
```

**Arguments**

|         |   |
|---------|---|
| df      | Data frame containing the elements to sample from.      |
| stratid | Variable in df containing the strata.                   |
| nsamp   | Variable in df containing the sample sizes.             |
| prn     | Variable in df containing the permanent random numbers. |

**Value**

Copy of the input data frame together with the boolean variable sampled.

**See Also**

[samp](#), [pps](#), [transformprn](#).

**Examples**

```
dfOut <- srs(df=ExampleData,
             nsamp="nsample",
             stratid="stratum",
             prn="rands")
```

---

|              |   |
|--------------|---|
| transformprn | <i>Permanent random number transformation</i> |
|--------------|---|

---

**Description**

Transformation of the permanent random numbers used in the sampling procedure, to control the overlap between samples, and thus control the sample coordination. The method used is specified in Lindblom and Teterukovsky (2007).

**Usage**

```
transformprn(df, prn, direction, start)
```

**Arguments**

|           |  |
|-----------|--|
| df        | Data frame containing the elements to sample from.   |
| prn       | Variable in df containing the permanent random numbers.  |
| direction | "U" or "R" for upwards, or to the right on the real-number line. "D" or "L" for downwards, or to the left on the real-number line. |
| start     | Starting point for the transformation.   |

**Value**

Copy of the input data frame together with the numeric variable `prn.old` containing the non-transformed permanent random numbers, as well as the numeric variable `prn` containing the transformed permanent random numbers.

**References**

Lindblom, A. and Teterukovsky, A. (2007) *Coordination of Stratified Pareto pps Samples and Stratified Simple Random Samples at Statistics Sweden*.

**See Also**

[samp](#), [srs](#), [pps](#).

**Examples**

```
dfOut <- transformprn(df=ExampleData, prn="rands", direction="U", start=0.2)
```

# Index

\*Topic **datasets**

ExampleData, [2](#)

ExampleData, [2](#)

pps, [2](#), [4](#), [5](#)

samp, [3](#), [3](#), [4](#), [5](#)

srs, [3](#), [4](#), [4](#), [5](#)

transformprn, [3](#), [4](#), [5](#)