Beautiful graphs with npde 3.0

Février 2021

Emmanuelle Comets, Romain Leroux

Contributors to npde: Karl Brendel, Marc Cerou, Romain Leroux, Thi Huyen Tram Nguyen, France Mentré

INSERM, IAME UMR 1137, Paris, France

Université de Paris, Paris, France.

npde website: www.npde.biostat.fr

```
[1] "name group is empty"
[1] "name predictor is empty"
[1] "name response is empty"
[1] "name covariates is empty"
[1] "name cens is empty"
[1] "name miss is empty"
[1] "name ipred is empty"
```

1 Running npde

Describe: - viral load data example - 20%

```
Viral load data with 20% censored data
[1] "name covariates is empty"
[1] "name miss is empty"
[1] "name ipred is empty"
---------------------------------------------
Distribution of npde :
      nb of obs: 300
           mean= 0.03715   (SE= 0.058 )
       variance= 0.9924   (SE= 0.081 )
       skewness= 0.1229
       kurtosis= 0.02054
---------------------------------------------

Statistical tests
  t-test                     : 0.519
  Fisher variance test       : 0.947
  SW test of normality       : 0.348
Global adjusted p-value      : 1
---
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
---------------------------------------------
Warfarin PK data
      npde for the model without covariates
[1] "name cens is empty"
[1] "name miss is empty"
[1] "name ipred is empty"
---------------------------------------------
Distribution of npde :
      nb of obs: 247
           mean= 0.03419   (SE= 0.06 )
       variance= 0.8753   (SE= 0.079 )
       skewness= -0.1149
       kurtosis= -0.0497
---------------------------------------------

Statistical tests
  t-test                     : 0.566
  Fisher variance test       : 0.157
  SW test of normality       : 0.371
Global adjusted p-value      : 0.471
---
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
--------------------------------------------
      npde for the model with covariates
[1] "name cens is empty"
[1] "name miss is empty"
[1] "name ipred is empty"
--------------------------------------------
Distribution of npde :
      nb of obs: 247
           mean= 0.02928   (SE= 0.059 )
       variance= 0.8549   (SE= 0.077 )
       skewness= -0.07211
       kurtosis= -0.4172
--------------------------------------------

Statistical tests
  t-test                     : 0.619
  Fisher variance test       : 0.096 .
  SW test of normality       : 0.368
Global adjusted p-value      : 0.288
---
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
--------------------------------------------
```

The `npde` package has built-in diagnostic plots which can be accessed for an object `y` returned by `autonpde()` (or `npde()` for the interactive execution) by the usual R command `plot()`:

```
plot(y)
```

```
> plot(yvir20)
```

produces the plot shown in figure **??**.

Eco TODO problems with the viral load data plots:

- not the same result when I execute interactively the viral load data (the p-value is NS interactively, but sometimes comes out significant)
- check the nubmer of simulations, maybe not enough ?

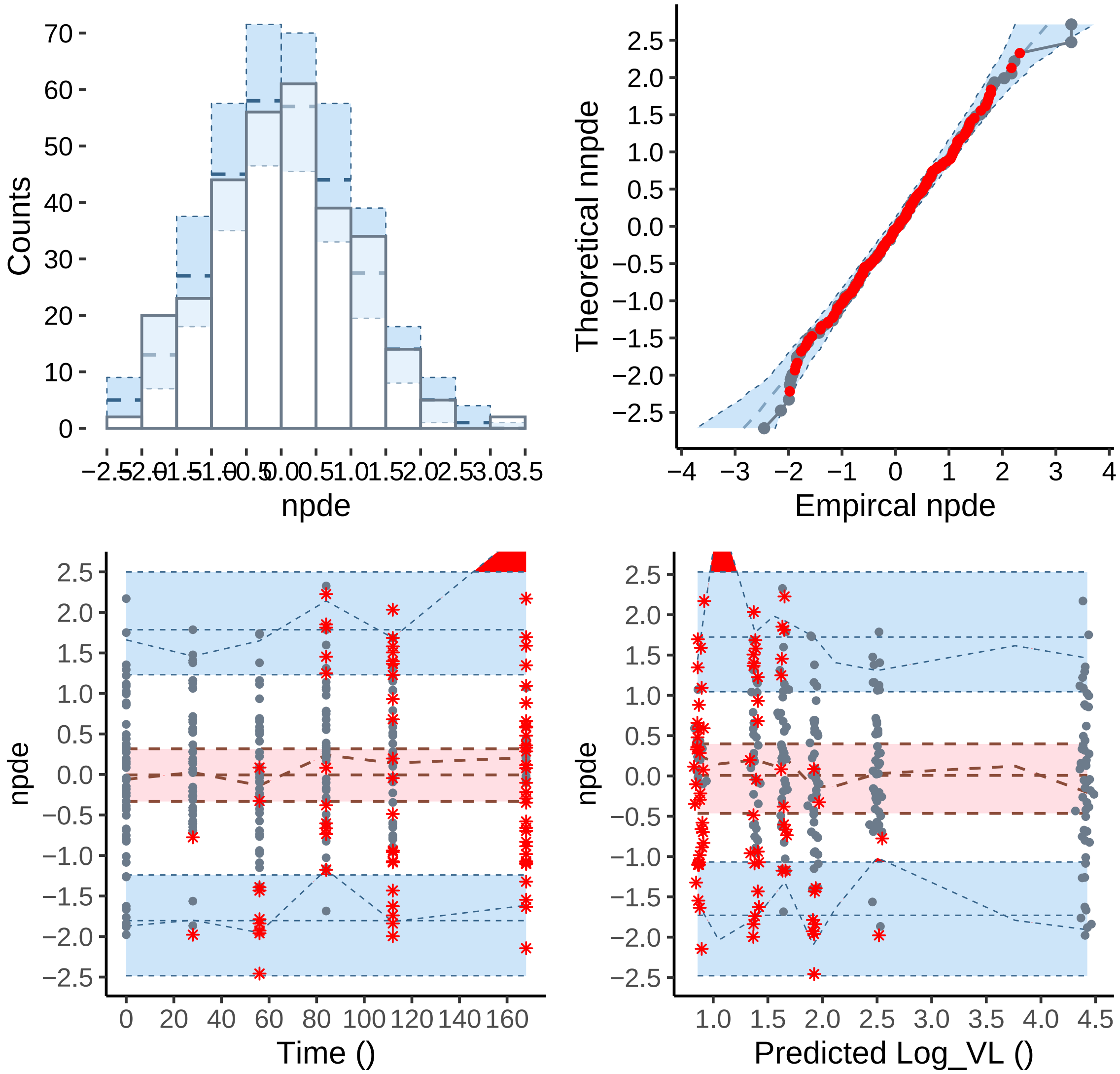

Figure 1: Default plots for the viral load data.

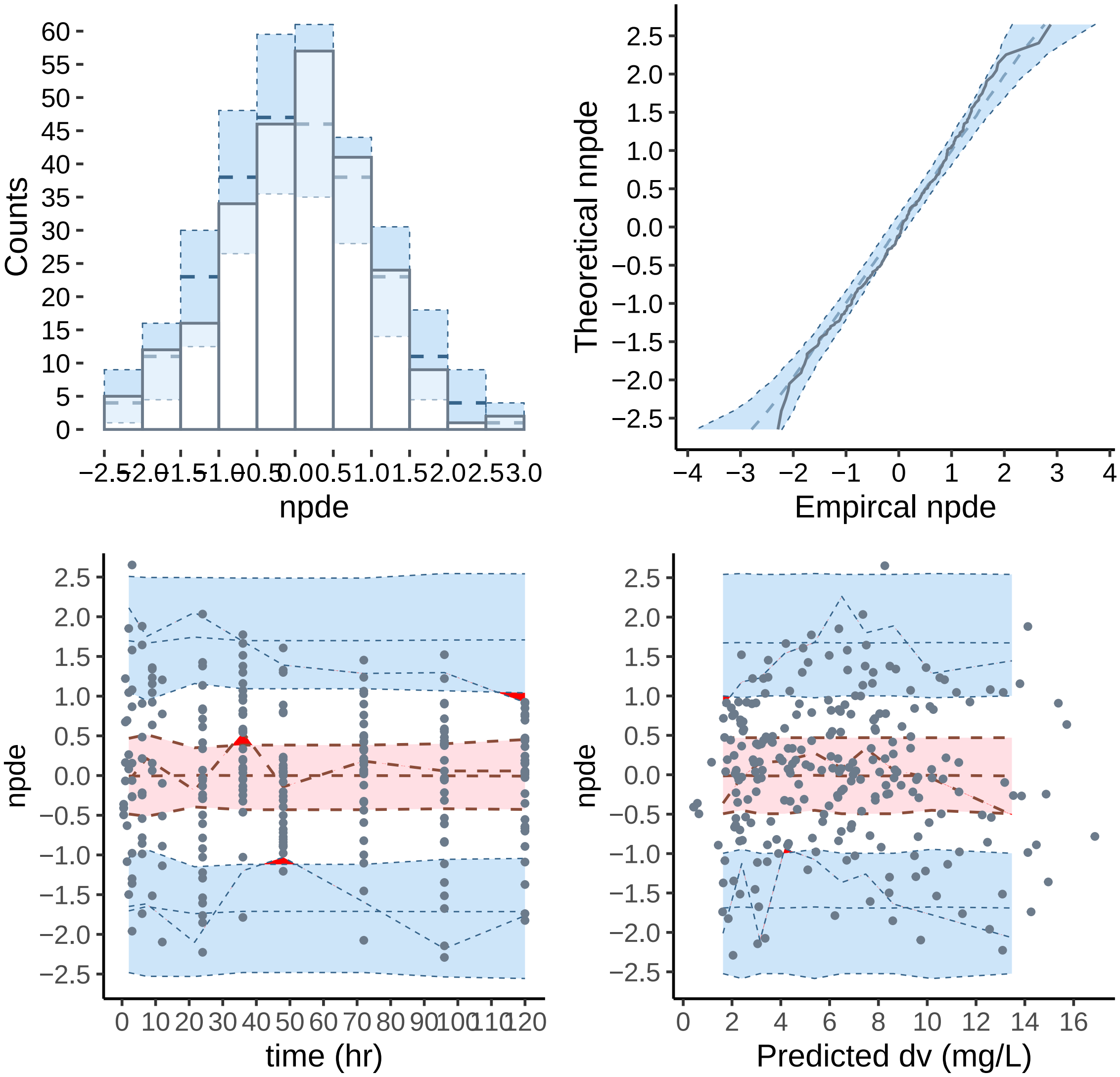

Figure 2: Default plots for the warfarin PK data, assuming a model without covariates.

2 Graphical options

2.1 Types of plots

Each of the default diagnostic plots, as well as a number of additional plots not shown by default, can also be produced on its own, using the argument `plot.type="type"`. Table **??** lists the plots that can be created in this way. Apart from VPC and the plot of the probability of an observation being under the LOQ, all plots can be obtained for npde, npdor pd(this is controlled by the `which=""` argument, defaulting to `which="npde"`).

Plot type	Description
data	Plots the observed data in the dataset
hist	Histogram of the npde
qqplot	QQ-plot of the npdeversus its theoretical distribution
ecdf	Empirical distribution function of the npde
x.scatter	Scatterplot of the npde versus the predictor X
pred.scatter	Scatterplot of the npde versus the population predicted values
cov.scatter	Scatterplot of the npde versus covariates
vpc	Plots a Visual Predictive Check (VPC)
loq	Plots the probability for an observation to be BQL, versus the predictor X

Table 1: Plot types available in the `npde` library. QQ-plots, histograms, cumulative cdf, and scatter plots can be produced for npde, pd or npd.

Table 2: tab:plottypes

To produce only the QQ-plot as a standalone graph, we request a ”qqplot”:

The plots can also be produced by the other metrics computed by `npde()`, for instance we might want to consider the histogram of pdwith the following command:

```
> plot(yvir20, plot.type="qqplot")
```

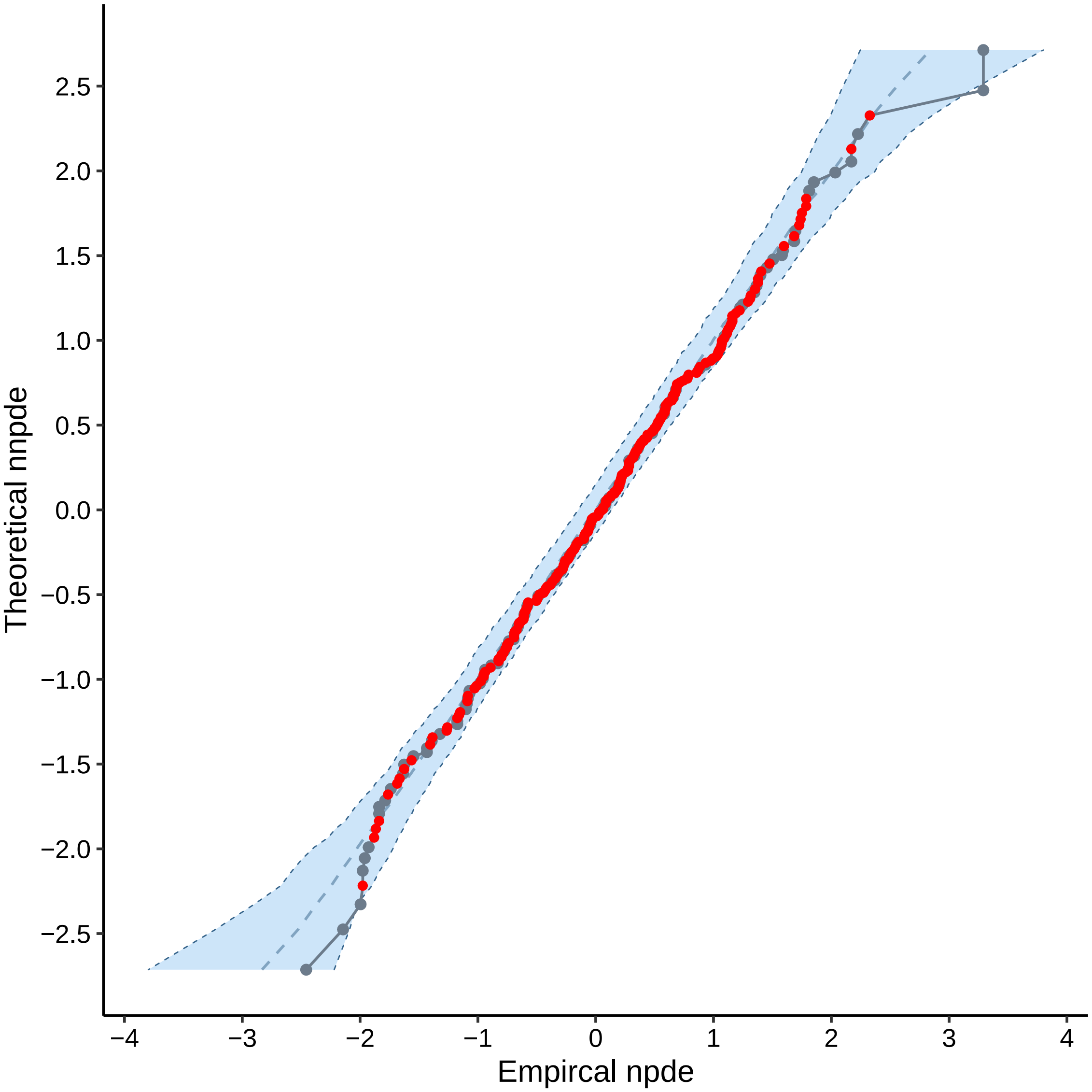

Figure 3: QQ-plot for the viral load data.

```
> plot(yvir20, plot.type="hist",which="pd")
```

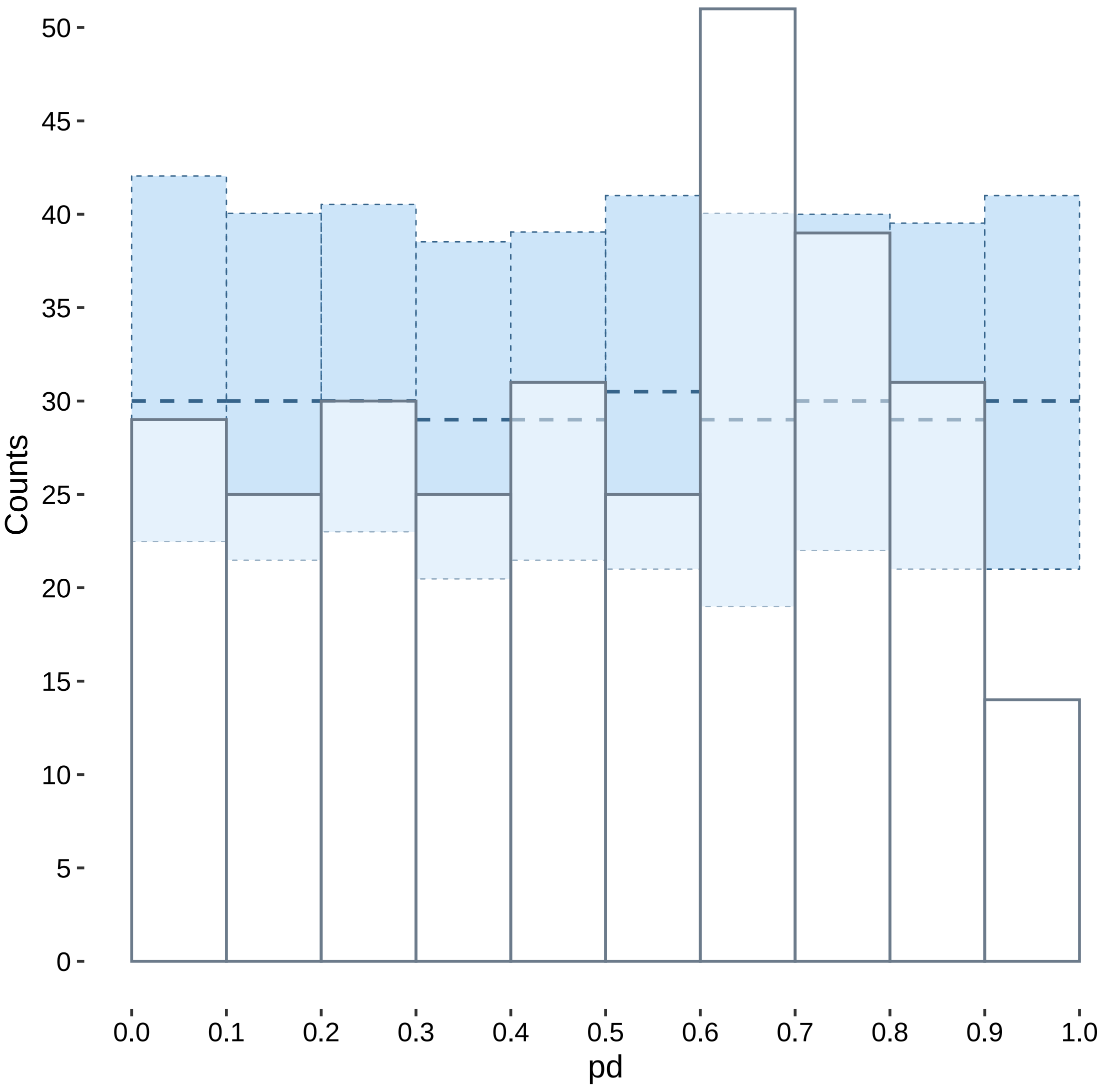

Figure 4: Histogram for the pd, computed in the viral load data example.

2.2 Changing graphical parameters

TODO: show different features in group, eg colours/symbols, titles/axes, boxes/grids
Decide on following options: verbose

Argument	Description	Default value
`verbose`	Output is produced for some plots (most notably when binning is used, this prints out the boundaries of the binning intervals) if TRUE	FALSE
`main`	Title	depends on plot
`sub`	Subtitle	empty
`size.main`	Size of the main title	14
`size.sub`	Size of the title for covariate	12
`xlab`	Label for the X-axis	depends on plot
`ylab`	Label for the Y-axis	depends on plot
`size.xlab`	Size of the label for the X-axis	12
`size.ylab`	Size of the label for the Y-axis	12
`breaks.x`	Number of tick marks on the X-axis	10
`breaks.y`	Number of tick marks on the Y-axis	10
`size.x.text`	Size of tick marks and tick labels on the X-axis	10
`size.y.text`	Size of tick marks and tick labels on the Y-axis	10
`xlim`	Range of values on the X-axis	empty, adjusts to the data
`ylim`	Range of values on the Y-axis	empty, adjusts to the data
`xaxt`	A character whether to plot the X axis. Specifying ”n” suppresses plotting of the axis	”y”
`yaxt`	A character whether to plot the Y axis. Specifying ”n” suppresses plotting of the axis	”y”
`xlog`	Scale for the X-axis (TRUE: logarithmic scale)	FALSE
`ylog`	Scale for the Y-axis (TRUE: logarithmic scale)	FALSE

Table 3: Graphical parameters that can be passed on the plot function: titles and axes.

Argument	Description	Default value
`col`	Main colour for observed data (applied to lines and symbols pertaining to observations if no other option is given to supersede this value)	”slategray4”
`lty`	Line type for observed data	1
`lwd`	Line width for observed data	0.5
`pch`	Symbol used to plot observed data	20
`alpha`	Transparencyfor observed data	1
`size`	Symbol size to plot observed data	1
`fill`	Colour used to fill area elements related to observed data (such as histogram bars)	”white

Table 4: Graphical parameters that can be passed on the plot function: colours and symbols.

2.3 Saving to file

2.3.1 Using ggsave from ggplot

2.3.2 Handling transparency with postscript files

Postcript doesn't handle the transparency used in the npde plots made with the ggplot2 library. Using ggsave(), the prediction intervals will likely be missing from the output. A workaround is to output the files to PDF format. Another is to use the following code

TODO

2.4 Arranging individual plots

3 Covariate plots

3.1 Plots of npde versus covariates

3.2 Stratified plots

The different diagnostic plots produced in section **??** can be stratified using the `covsplit=TRUE` option. By default, the plots will be stratified for each covariate separately. The plots produced depend on the nature of the covariate, following [**?**]:

- for categorical covariates, a plot is created for each category of the covariate
- for continuous covariates, three plots

4 Reference profile

5 Waffle plot

```
[1] "name cens is empty"
[1] "name miss is empty"
[1] "name ipred is empty"
---------------------------------------------
Distribution of npde :
      nb of obs: 247
           mean= 0.03419   (SE= 0.06 )
       variance= 0.8753   (SE= 0.079 )
       skewness= -0.1149
       kurtosis= -0.0497
---------------------------------------------

Statistical tests
  t-test                    : 0.566
  Fisher variance test      : 0.157
  SW test of normality      : 0.371
Global adjusted p-value     : 0.471
---
Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
---------------------------------------------
```

```
> do.call(grid.arrange, c(p1, p2, p3, p1, p2, list( nrow=2, ncol=3)))
```

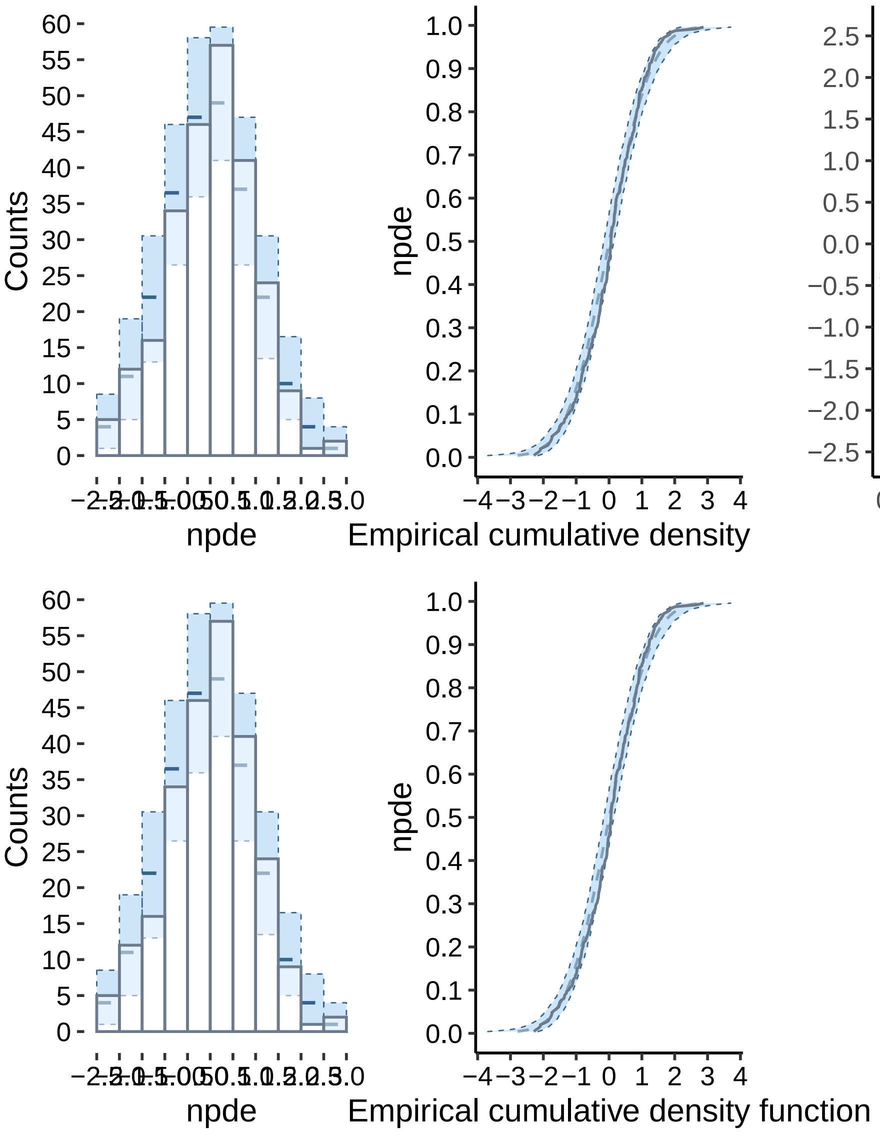

Figure 5: Waffle plot for the warfarine data.