**Versión en español.**

**Explicación del código del paquete SEROFOI**

**Las librerías**

Las siguientes son librerías necesarias para el paquete de serofoi, la descripción de su uso se encuentra a continuación.

**Rstan:** El paquete R rstan proporciona RStan, la interfaz R para Stan. El paquete rstan permite ajustar convenientemente los modelos Stan de R (R Core Team 2014) y acceder a la salida, incluidas las inferencias posteriores y las cantidades intermedias, como las evaluaciones de la densidad logarítmica posterior y sus gradientes.

**Tidyverse:** El tidyverse es una colección obstinada de paquetes R diseñados para la ciencia de datos. Todos los paquetes comparten una filosofía de diseño, gramática y estructuras de datos subyacentes. El tidyverse central incluye los paquetes que probablemente usará en los análisis de datos cotidianos.

**Reshape2:** Reshape2 es un reinicio del paquete de remodelación. Reshape2 usa ese conocimiento para crear un nuevo paquete para remodelar datos que es mucho más centrado y mucho más rápido. Existe una librería de tidyverse que puede reemplazarla.

**Bayesplot:** bayesplot es un paquete de R que proporciona una amplia biblioteca de funciones de trazado para su uso después de ajustar modelos bayesianos (normalmente con MCMC). Los gráficos creados por bayesplot son objetos ggplot, lo que significa que después de crear un gráfico, se puede personalizar aún más utilizando varias funciones del paquete ggplot2.

**Loo:** Validación cruzada eficiente aproximada de dejar uno fuera (LOO) para el ajuste de modelos bayesianos utilizando la cadena de Markov Monte Carlo. La aproximación utiliza el muestreo de importancia suavizado de Pareto (PSIS), un nuevo procedimiento para regularizar los pesos de importancia. Como subproducto de los cálculos, también obtenemos errores estándar aproximados para errores predictivos estimados y para la comparación de errores predictivos entre modelos. El paquete también proporciona métodos para utilizar el apilamiento y otras técnicas de ponderación de modelos para promediar las distribuciones predictivas bayesianas.

**Pracma:** Proporciona una gran cantidad de funciones de análisis numérico y álgebra lineal, optimización numérica, ecuaciones diferenciales, series de tiempo, además de algunas funciones matemáticas especiales conocidas. Utiliza nombres de funciones de matlab cuando corresponda para simplificar la transferencia.

**Cowplot:** El paquete cowplot proporciona varias funciones que ayudan a crear figuras con calidad de publicación, como un conjunto de temas, funciones para alinear gráficos y organizarlos en figuras compuestas complejas, y funciones que facilitan la anotación de gráficos o la combinación de gráficos con imágenes.

**Grid:** El paquete grid en R implementa las funciones gráficas primitivas que subyacen en el sistema de trazado ggplot2. El paquete grid proporciona un conjunto de funciones y clases que representan objetos gráficos o grobs, que se pueden manipular como cualquier otro objeto R.

**GridExtra:** Proporciona una serie de funciones a nivel de usuario para trabajar con gráficos de "cuadrícula", en particular para organizar múltiples gráficos basados en cuadrículas en una página y dibujar tablas.

**Hmisc:** La librería Hmisc contiene muchas funciones útiles para el análisis de datos, gráficos de alto nivel, operaciones de servicios públicos, funciones para calcular el tamaño y la potencia de la muestra, traducir conjuntos de datos SAS a R, imputar valores faltantes, creación avanzada de tablas, agrupamiento de variables, manipulación de cadenas de caracteres, conversión de objetos R a código LaTeX, las variables de codificación y el análisis de medidas repetidas de arranque.

**Dplyr:** Uno de los paquetes principales de tidyverse en el lenguaje de programación R, dplyr es principalmente un conjunto de funciones diseñadas para permitir la manipulación de marcos de datos de una manera intuitiva y fácil de usar. Se suele utilizar dplyr para transformar los conjuntos de datos existentes en un formato más adecuado para algún tipo particular de análisis o visualización de datos.

**Epitrix:** Una colección de pequeñas funciones útiles para el análisis de epidemias y el modelado de enfermedades infecciosas. Esto incluye el cálculo de los números básicos de reproducción a partir de las tasas de crecimiento, la generación de etiquetas hash para anonimizar los datos y el ajuste de distribuciones Gamma discretas.

**Gsubfn:** La función gsubfn es como gsub pero puede tomar una función de reemplazo u otros objetos en lugar de la cadena de reemplazo. Las coincidencias y las referencias anteriores se ingresan en la función de reemplazo y se reemplazan por la salida de la función. gsubfn se puede usar para dividir cadenas según el contenido en lugar de los delimitadores y para la interpolación de cadenas al estilo cuasi-perl. El paquete también tiene facilidades para traducir fórmulas a funciones y permitir tales fórmulas en llamadas a funciones en lugar de funciones. Esto se puede usar con funciones de R como apply, sapply, lapply, optim, integra, xyplot, Filter y cualquier otra función que espere otra función como argumento de entrada o funciones como llamadas cat o sql que pueden involucrar cadenas donde la sustitución es deseable. También hay una función para devolver múltiples objetos de funciones y una versión de transform que permite que RHS se refiera a LHS utilizado en la misma transform.

**vscDebbuger:** El depurador funciona de la siguiente manera: se inicia un proceso R dentro de un proceso secundario. Se carga el paquete R vscDebugger. El depurador inicia y controla los programas R enviando información a la entrada estándar del proceso secundario.

**Funciones finales**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Nombre** | **Agr. código** | **Agr. Esquema** | **Descripción** | **Comentarios** |
| **1** | check\_all\_diagnostics | fUtility |  | Revisa que todo esté funcionando correctamente en el fit. |  |
| **2** | cleanObject | fUtility |  | Transmite el archivo stan fit. |  |
| **3** | compare\_and\_save\_best\_model | fCheck |  | Compara y guarda el mejor modelo. |  |
| **4** | dir\_results | fUtility |  | Guarda los resultados en un lugar específico. |  |
| **5** | extract\_and\_save | fCheck |  | Extrae y guarda. |  |
| **6** | extract\_summary\_mod | fCheck |  | Extrae un resumen de los modelos. |  |
| **7** | fCombinedPlots | fPlot |  | Combina los plots en uno solo. |  |
| **8** | fFitModel | fFitting |  | Ajusta los modelos. |  |
| **9** | fFitModel\_log | fFitting |  | Ajusta el modelo log. |  |
| **10** | get\_exposure\_matrix | fUtility |  | Obtiene la matriz de exposición. |  |
| **11** | get\_model\_comparison\_plot | fPlot |  | Obtiene la gráfica para comparar los modelos. |  |
| **12** | get\_posterior\_summary | fUtility |  | Consigue un resumen posterior de los resultados. |  |
| **13** | get\_prev\_expanded | fUtility |  |  |  |
| **14** | get\_residuals | fUtility |  | Obtiene los residuos. |  |
| **15** | get\_table\_rhats | fCheck |  | Obtiene la tabla de convergencia de los modelos. |  |
| **16** | make\_yexpo | fUtility |  |  |  |
| **17** | obtain\_prevalence\_extended | fUtility |  |  |  |
| **18** | partition\_div | fUtility |  | Devuelve arreglos de parámetros separados en transiciones divergentes y no divergentes. |  |
| **19** | plot\_info\_table | fPlot |  | Plotea la tabla con información. |  |
| **20** | RunSaveModels | fRun |  | Corre los modelos considerados. |  |
| **21** | sub\_sample | fUtility |  | Obtiene una muestra. |  |
| **22** | thin\_chain | fUtility |  |  |  |
| **23** | vertical\_plot\_arrange\_per\_model | fPlot |  |  |  |
| **24** | get\_data\_binned | fUtility |  | Obtiene datos agrupados. | No necesaria y su eliminación reduce el tiempo de ejecución. |
| **25** | Constant\_foi\_Bi | Stanmodels | declare\_design  declare\_model | Se define el modelo para una fuerza de infección constante con un prior dado por una distribución uniforme con parámetros 0 y 2. |  |
| **26** | Continuous\_foi\_normal\_Bi | Stanmodels | declare\_design  declare\_model | Se define el modelo para una fuerza de infección variable de manera anual con prior y upper prior dados por distribuciones normales y de Pareto y Cauchy respectivamente. |  |
| **27** | Continuous\_foi\_normal\_log | Stanmodels | declare\_design  declare\_model | Define el modelo para una fuerza de infección variable anual con un prior de baja transmisión dada por el logaritmo de la fuerza de infección. Los prior están dados por distribuciones normales, mientras el upper prior viene dado por una distribución de Cauchy. |  |

**English versión**

**Explanation of the SEROFOI package code**

**The packages**

The following are required libraries for the serofoi package, the description of their use is given below.

**Rstan:** The R package rstan provides RStan, the R interface to Stan. The rstan package allows you to conveniently fit R Stan models (R Core Team 2014) and access the output, including posterior inferences and intermediate quantities, such as evaluations of the posterior logarithmic density and its gradients.

**Tidyverse:** The tidyverse is a opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures. The core tidyverse includes the packages you're likely to use in your day-to-day data analysis.

**Reshape2:** Reshape2 is a reshape pack reboot. Reshape2 uses that knowledge to create a new data reshaping package that is much more focused and much faster. There is a tidyverse library that can replace it.

**Bayesplot:** Bayesplot is an R package that provides an extensive library of plotting functions for use after fitting Bayesian models (usually with MCMC). The plots created by bayesplot are ggplot objects, which means that after a plot is created, it can be further customized using various functions of the ggplot2 package.

**Loo:** Approximate efficient leave-one-out (LOO) cross-validation for Bayesian model fitting using Markov chain Monte Carlo. The approximation uses Pareto smoothed importance sampling (PSIS), a new procedure to regularize importance weights. As a byproduct of the calculations, we also obtain approximate standard errors for estimated predictive errors and for the comparison of predictive errors between models. The package also provides methods for using stacking and other model weighting techniques to average Bayesian predictive distributions.

**Pracma:** Provides a large number of functions for numerical analysis and linear algebra, numerical optimization, differential equations, time series, in addition to some well-known special mathematical functions. Use matlab function names where appropriate to simplify transfer.

**Cowplot:** The cowplot package provides several features that help create publication-quality figures, such as a set of themes, features for aligning graphics and organizing them into complex composite figures, and features that make it easy to annotate graphics or combine graphics with images. .

**Grid:** The grid package in R implements the primitive graphics functions that underlie the ggplot2 plotting system. The grid package provides a set of functions and classes that represent graphical objects, or grobs, that can be manipulated like any other R object.

**GridExtra:** Provides a number of user-level functions for working with "grid" charts, in particular for arranging multiple grid-based charts on a page and drawing tables.

**Hmisc:** The Hmisc library contains many useful functions for data analysis, high-level graphing, utility operations, functions for calculating sample size and power, translating SAS datasets to R, imputing missing values, creating advanced table mapping, variable grouping, string manipulation, conversion of R objects to LaTeX code, encoding variables, and bootstrap repeated measures analysis.

**Dplyr:** One of the tidyverse's main packages in the R programming language, dplyr is primarily a set of functions designed to allow manipulation of data frames in an intuitive and user-friendly way. dplyr is often used to transform existing datasets into a format more suitable for some particular type of data analysis or visualization.

**Epitrix:** A collection of small functions useful for epidemic analysis and infectious disease modeling. This includes calculating basic reproduction numbers from growth rates, generating hash tags to anonymize the data, and fitting discrete Gamma distributions.

**Gsubfn:** The gsubfn function is like gsub but it can take a replacement function or other objects instead of the replacement string. Matches and previous references are fed into the replace function and replaced by the output of the function. gsubfn can be used to split strings based on content rather than delimiters and for quasi-perl-style string interpolation. The package also has facilities to translate formulas into functions and allow such formulas in function calls instead of functions. This can be used with R functions like apply, sapply, lapply, optim, integra, xyplot, Filter and any other function that expects another function as an input argument or functions like cat or sql calls that can involve strings where substitution is desirable . There is also a function to return multiple function objects and a version of transform that allows RHS to refer to LHS used in the same transform.

**vscDebugger:** The debugger works as follows: an R process is started within a child process. The vscDebugger R package is loaded. The debugger starts and controls R programs by sending information to the standard input of the child process.

**Final functions**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Name** | **Gr. Code** | **Gr. schema** | **Description** | **Comments** |
| **1** | check\_all\_diagnostics | fUtility |  |  |  |
| **2** | cleanObject | fUtility |  |  |  |
| **3** | compare\_and\_save\_best\_model | fCheck |  |  |  |
| **4** | dir\_results | fUtility |  |  |  |
| **5** | extract\_and\_save | fCheck |  |  |  |
| **6** | extract\_summary\_mod | fCheck |  |  |  |
| **7** | fCombinedPlots | fPlot |  |  |  |
| **8** | fFitModel | fFitting |  |  |  |
| **9** | fFitModel\_log | fFitting |  |  |  |
| **10** | get\_exposure\_matrix | fUtility |  |  |  |
| **11** | get\_model\_comparison\_plot | fPlot |  |  |  |
| **12** | get\_posterior\_summary | fUtility |  |  |  |
| **13** | get\_prev\_expanded | fUtility |  |  |  |
| **14** | get\_residuals | fUtility |  |  |  |
| **15** | get\_table\_rhats | fCheck |  |  |  |
| **16** | make\_yexpo | fUtility |  |  |  |
| **17** | obtain\_prevalence\_extended | fUtility |  |  |  |
| **18** | partition\_div | fUtility |  |  |  |
| **19** | plot\_info\_table | fPlot |  |  |  |
| **20** | RunSaveModels | fRun |  |  |  |
| **21** | sub\_sample | fUtility |  |  |  |
| **22** | thin\_chain | fUtility |  |  |  |
| **23** | vertical\_plot\_arrange\_per\_model | fPlot |  |  |  |
| **24** | get\_data\_binned | fUtility |  | get grouped data | It is not necessary and its removal reduces execution time. |
| **25** | Constant\_foi\_Bi | Stanmodels | declare\_design  declare\_model | The model is defined for a constant infection force with a prior given by a uniform distribution with parameters 0 and 2. |  |
| **26** | Continuous\_foi\_normal\_Bi | Stanmodels | declare\_design  declare\_model | The model is defined for an annual variable infection force with prior and upper prior given by normal and Pareto and Cauchy distributions, respectively. |  |
| **27** | Continuous\_foi\_normal\_log | Stanmodels | declare\_design  declare\_model | Define the model for an annual variable infection force with a low transmission prior given by the logarithm of the infection force. The priors are given by normal distributions, while the upper prior is given by a Cauchy distribution. |  |

**Referencias**

1. <https://www.rdocumentation.org/packages/>
2. <https://cran.r-project.org/web/packages/>