

# Libro dinámico y colaborativo de Estadística descriptiva

Ángel Berihuete Macías\*, Carmen Dolores Ramos García, Juan Antonio García Ramos

\*Departamento de Estadística e Investigación Operativa, Facultad de Ingeniería Naval y Oceánica

[angel.berihuete@uca.es](mailto:angel.berihuete@uca.es)

**RESUMEN:** Existe una apuesta clara por parte del Servicio de Publicaciones de la Universidad de Cádiz hacia la autoedición de materiales docentes por parte de los profesores. Este hecho se constata claramente con la nueva colección de manuales de Matemáticas y Física, donde los profesores tendrán que utilizar una plantilla determinada en la plataforma Overleaf para generar su manual. El presente proyecto propone aprovechar un nuevo paradigma en la escritura de libros de contenido estadístico a través de la participación directa en la edición de los contenidos por parte del docente. Dicha participación se realiza a través de herramientas especializadas de edición con programas de ámbito matemático/estadístico, pudiendo obtener, como resultado final, un libro con contenidos dinámicos, así como su publicación en distintos formatos digitales.

**PALABRAS CLAVE:** Autoedición, proyecto colaborativo, Estadística, lenguaje markdown, Rstudio

## INTRODUCCIÓN

El hecho de publicar un manual de Estadística descriptiva o inferencial que sirva de ayuda a los estudiantes de Grado conlleva una serie de retos:

a) No es sencillo escribir formulación matemática de calidad con editores de texto como Word del paquete Microsoft Office, o Writer del paquete Open Office. Este tipo de editores, también llamados WYSIWYG en sus siglas en inglés necesitan de complementos específicos para insertar contenido estadístico y/o fórmulas.

b) Los libros de Estadística no incluyen en la actualidad contenido dinámico con el que el estudiante pueda interactuar.

c) La actualización de un manual de Estadística es tediosa y lenta, pudiéndose tardar varios años en volver a sacar una nueva edición en la que se corrigen errores de ediciones anteriores y se actualiza el material obsoleto (versiones de software, etc.)

d) La revisión de un manual de Estadística recae generalmente en dos o tres revisores, y una vez que está editado, no pueden añadirse modificaciones hasta su próxima revisión.

En los últimos cinco años han emergido herramientas digitales que ayudan a solventar los problemas descritos anteriormente. Las herramientas tienen un carácter más general, pero son de fácil aplicación a manuales estadísticos orientados a estudiantes de Grado.

En las siguientes secciones describimos nuestra experiencia utilizando dichas herramientas, analizando los pros y los contras cuando se actualizan/crean los materiales del aula a formatos interactivos para el estudiante.

## PUESTA EN MARCHA DE LOS SERVICIOS E INSTALACIÓN DEL SOFTWARE NECESARIO

Antes de abordar la edición de material interactivo y colaborativo para los estudiantes, debemos poner en marcha los servicios necesarios para optimizar el flujo del trabajo en la creación de contenidos. Hemos creado una cuenta y un repositorio específico para este proyecto en la plataforma

[GitHub](#) (1). Elegimos esta plataforma por ofrecer planes gratuitos de almacenaje con repositorios privados a la comunidad universitaria.

GitHub es una plataforma de desarrollo colaborativo de software que puede almacenar proyectos utilizando un sistema de control de versiones (Git). Los proyectos pueden almacenarse en *repositorios* privados o públicos. En el caso de este proyecto optamos por un repositorio público, pudiendo tener acceso todos los usuarios. La fortaleza de GitHub frente a servicios como Dropbox u Owncloud, radica en que se asienta en el control de versiones Git. Además tiene herramientas que hacen que la propuesta de corrección de los contenidos sea fácil para el usuario final.

Hemos creado una cuenta gratuita en la plataforma [RStudio Connect](#) (2) para la publicación del material interactivo. En realidad esta cuenta gratuita sirve para testear el contenido creado, pero se necesitará un plan de pago específico si se quieren ofrecer los servicios de descarga o visualización de contenido a mayor escala. Como alternativa a los planes de pago en la plataforma RStudio Connect, hemos utilizado el servicio *pages* de GitHub para publicar nuestro material.

Finalmente hemos instalado el software necesario para la creación de material interactivo, principalmente [RStudio](#) (3), [Bookdown](#) (4) y [Shiny](#) (5). Estas herramientas hacen que la autoedición de material interactivo Estadístico sea rápido y eficaz. En el proceso de instalación de las herramientas anteriores, se instalaron de forma automática gran cantidad de paquetes de [R](#) (6) y hubo que instalar librerías necesarias para que dichos paquetes funcionaran. Destacar entre ellas la librería [Pandoc](#) (7) que se utiliza como conversor de documentos de diferentes formatos.

Todos los programas y herramientas instaladas, así como el sistema operativo que se ha utilizado en todo el proyecto, se basa en software libre.

## ACTUALIZACIÓN DEL MATERIAL UTILIZADO EN EL AULA

Hemos centrado la actualización del material en las presentaciones que realizamos en clase, generalmente clases

magistrales en las que se proyecta una presentación del tipo PowerPoint. En nuestro caso, dicho material se realizó con plantillas [Beamer de LaTeX](#) (8), un lenguaje de etiquetado que necesita ser compilado para obtener un texto con expresiones matemáticas de gran calidad. Aunque ya existen en el mercado plataformas y programas que hacen más sencilla la edición en lenguaje LaTeX, en general, obtener un texto que incluya gráficos, tablas y ecuaciones matemáticas es complejo. Además, en dichas presentaciones los elementos gráficos suelen ser imágenes estáticas insertadas en el texto a partir de archivos en formato PNG o PDF.

Puesto que nuestro objetivo final es crear material interactivo mediante Rstudio, Bookdown y Shiny, sustuiremos el formato LaTeX a lenguaje [markdown](#) (9) para una edición sencilla evitando la mayoría del etiquetado que necesita LaTeX. Es importante recalcar que el lenguaje markdown tiene su propio etiquetado, pero es más sencillo que el lenguaje LaTeX. En el siguiente enlace se explica cuál fue nuestra rutina a la hora de cambiar de un lenguaje a otro:

<https://gist.github.com/AngelBerihuete/a2fbcf85a901eaeeddda69655b6fc82c>

Los gráficos y tablas que aparecen en la presentación del aula se crean en el nuevo proyecto de manera directa a través de trozos de código de R. Esto se hace de manera muy sencilla utilizando el programa RStudio. Estos trozos, llamados *chunks*, hacen que cada vez que se compile de nuevo el material, los gráficos y tablas se generan de nuevo, haciendo que los contenidos tengan la última versión en todo momento. Desde el punto de vista estadístico esto es muy importante ya que, cambiando los datos, o los parámetros de un modelo, volvemos a obtener un material completamente actualizado habiendo hecho pequeñas modificaciones en el texto. Puede encontrarse más información sobre la composición de texto matemático, resultados y gráficos estadísticos en la página web de [rmarkdown](#) (10).

Finalmente, varios de los gráficos que se incluyen en las presentaciones han sido transformadas a contenido dinámico con Shiny en incluidas en el nuevo material para que el estudiante pueda interactuar con ellas. Las aplicaciones Shiny son aplicaciones web *open source* que utilizan el lenguaje estadístico R. Shiny te ayuda a realizar análisis interactivos en la web sin necesidad de conocer lenguajes específicos como HTML, CSS o JavaScript. Su página web contiene multitud de recursos para los profesores e investigadores, incluido un plan de soporte y planes de pago para almacenar las aplicaciones que se generen.

## RESULTADOS

Puede verse el primer borrador del material interactivo en la página web de Github (11). La página web muestra de manera automática una barra lateral para navegar por los contenidos, así como un menú superior para una búsqueda rápida de contenidos, ajuste del tamaño de las fuentes, edición del contenido en Github y descarga del libro en formato PDF y EPUB. También aparecen dos iconos con los que puede ir directamente al proyecto en la plataforma Github o compartir el proyecto en las redes sociales, ver Figura 1.

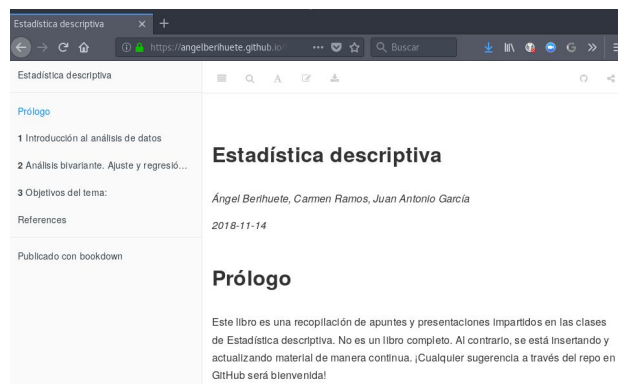


Figura 1. Imagen de la página web creada para el proyecto.

Tiene especial relevancia en nuestro proyecto el enlace directo para editar el contenido de alguna de las secciones. Si el lector tiene una cuenta en la plataforma Github, podrá sugerir la corrección o incorporación de contenido mediante un *pull-request*, es decir, cualquier estudiante o profesor podrá pedir incorporar su código en R, fórmulas y/o ejemplos aclaratorios del texto en cuestión. Los autores revisarán dicha petición y si están conforme con los cambios harán una *merge* con las sugerencias del lector actualizando de esa manera el libro, ver Figura 2.

El contenido interactivo es embebido desde la plataforma Shiny en el libro, pudiendo interactuar con él en formato HTML. Esta opción no está disponible para formato PDF o EPUB. En estos formatos aparece en su lugar una imagen estática que representa dicho contenido.

El libro está en la actualidad en su fase preliminar, no pudiéndose entregar aún a los estudiantes:

<https://angelberihuete.github.io/EstNavBook/index.html>

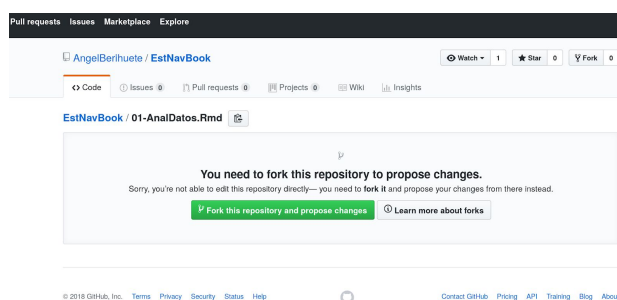


Figura 2. Imagen de un pull-request al contenido del proyecto.

## DISCUSIÓN Y CONCLUSIONES

El material creado no difiere en conceptos o procedimientos del material que utilizado en el aula dentro de la asignatura. Sin embargo, el hecho de ser un material interactivo hace que los estudiantes profundicen más en los conceptos, haciéndose preguntas del tipo *¿qué pasaría si ...?* cuando interactúan con un elemento creado con Shiny. Por ejemplo, un estudiante puede preguntarse qué pasaría si cambio la longitud de los intervalos en un histograma, o qué pasaría si cambio los parámetros de un modelo probabilístico

como la distribución normal, pudiendo comprobar en ese mismo momento el resultado a su pregunta.

Este tipo de preguntas, y la respuesta inmediata del material interactivo, hace que optimicemos las clases prácticas de Estadística, centrándonos más en los procedimientos y no repasando los conceptos. Sin embargo, la edición y/o corrección de los contenidos por parte de los estudiantes no ha sido satisfactorio. El material está pensado para estudiantes de primer curso del Grado, y estos estudiantes no tienen la capacidad de aprender las herramientas necesarias y hacer una edición adecuada de los materiales en su primer año de Universidad. Creemos que la mejora de los materiales debe llevarse a cabo por profesores del mismo Departamento (o profesores de otras Universidades) previo curso introductorio sobre el buen uso de las herramientas involucradas.

Más aún, las plataformas de almacenaje y publicación de los materiales son externas a la Universidad de Cádiz. Esto supone que debemos poner especial cuidado en el tipo de licencia con el que publicaremos nuestros materiales. En el caso de este proyecto queríamos que todos los materiales fueran de acceso libre, de forma que otros profesores y/o estudiantes pudieran reeditar y/o aportar correcciones al mismo, de manera que optamos por una licencia de Creative Commons Reconocimiento CompartirIgual 4.0 Internacional.

## AGRADECIMIENTOS

Agradecemos al Vicerrectorado de Recursos Docentes y de la Comunicación de la Universidad de Cádiz la financiación de este proyecto en en la convocatoria de Proyectos de Innovación y Mejora Docente (INNOVA) del curso 2017/2018.

## REFERENCIAS

1. Plataforma GitHub. <https://github.com/>
2. Plataforma RStudio Connect.  
<https://beta.rstudioconnect.com/connect/#/welcome>
3. RStudio: Integrated Development Environment for R, RStudio Team, RStudio, Inc., Boston, MA, 2016, <http://www.rstudio.com/>
4. Paquete Bookdown. <https://bookdown.org/>
5. Paquete Shiny. <https://shiny.rstudio.com/>
6. R: A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, 2018, <https://www.R-project.org/>
7. Librería Pandoc. <https://pandoc.org/>
8. El paquete Beamer de LaTeX.  
<https://ctan.org/pkg/beamer?lang=en>
9. Lenguaje markdown.  
<https://en.wikipedia.org/wiki/Markdown>
10. rmarkdown: Dynamic Documents for R, JJ Allaire *et al.*, 2018, R package version 1.10,  
<https://CRAN.R-project.org/package=rmarkdown>
11. Página web que muestra los contenidos del proyecto.  
<https://angelberihuete.github.io/EstNavBook/index.html>