

Mejorando la comprensión de conceptos estadísticos mediante aplicaciones interactivas innovadoras

Introducción

La Estadística es una materia obligatoria en los grados de Ciencias e Ingeniería, así como en muchos grados de Ciencias Sociales. Así, es un campo de conocimiento transversal que se aplica a una multitud de problemas de diversa índole, pero que se pueden abordar con la misma técnica estadística.

La experiencia nos dice que, en muchas ocasiones, la asignatura no despierta suficiente interés entre el estudiantado. Muchas pueden ser las causas de este desinterés, entre otras las siguientes: (1) Encontrarse en primer curso, donde todavía no ven la aplicabilidad; (2) Falta de base matemática para abordar ciertos problemas; (3) Predominio de metodologías docentes clásicas frente a otras más innovadoras; (4) Predominio de las explicaciones teóricas frente al trabajo empírico; (5) Uso de ejemplos y casos de estudio genéricos o lejos de su ámbito de conocimiento.

Pero independientemente de los motivos, nos gustaría destacar en este trabajo la dificultad para entender conceptos abstractos que tienen los estudiantes. Y el hecho de que esta dificultad se extiende más allá de los estudios de grado, encontrándonos egresados en distintos sectores de actividad que no son capaces de afrontar con solvencia análisis estadísticos que se supone aprendieron en su etapa en la universidad.

En el Grupo de Innovación Docente “Laboratorio de Ciencia de Datos para la Innovación de la Enseñanza” trabajamos para facilitar el uso de técnicas y métodos de Ciencia de Datos a la innovación de la enseñanza. En este trabajo proponemos el uso de aplicaciones interactivas innovadoras que ayuden, por un lado, a la comprensión de los conceptos estadísticos, y por otro a la mayor implicación y motivación del alumnado en las asignaturas.

Material y métodos

Para el desarrollo de las aplicaciones interactivas, utilizamos el software estadístico y lenguaje de programación R (<https://www.r-project.org>). R es desde hace años el estándar “de facto” para análisis de datos. No obstante, al ser también lenguaje de programación, sus posibilidades son ilimitadas, como demuestran los más de 18.000 paquetes desarrollados por la comunidad para una gran variedad de dominios de aplicación. Algunos paquetes constituyen auténticos “frameworks” para el desarrollo de aplicaciones, como es el caso del paquete “shiny” (<https://shiny.rstudio.com>). Con shiny podemos desarrollar aplicaciones interactivas que se ejecutan en el navegador, sin necesidad de saber programar páginas web. Todo se genera mediante funciones de R de forma análoga a como se haría un análisis de datos.

En este trabajo presentamos una aplicación para ilustrar la definición de probabilidad como frecuencia relativa en el límite. Un ejemplo clásico sería lanzar una moneda al aire repetidas veces y contar el número de caras que salen en una sucesión de n “experimentos”. La probabilidad de obtener cara sería, asintóticamente, el número de caras dividido entre el total de tiradas. Una actividad “manipulativa” podría consistir en poner a los estudiantes a lanzar monedas, anotar las caras y hacer los cálculos. Pero llevaría demasiado tiempo para un concepto básico.

Los métodos y herramientas de Ciencia de Datos, como la simulación y el software estadístico R, nos permiten simular cualquier experimento aleatorio, un número grande de veces, y personalizando para cada dominio de aplicación.

Resultados

Hemos desarrollado una aplicación interactiva para explicar este importante concepto, que está publicada en https://elcano.shinyapps.io/probability_as_relative_frequency/. La ventaja frente al ejemplo de la moneda, es que podemos utilizar “sucesos” más próximos a los intereses de los estudiantes. Por ejemplo, en el grado en Videojuegos podemos hablar de la probabilidad de que aparezca un atacante o un defensor; en el grado en Biología, podemos hablar de la probabilidad de que un tratamiento sea efectivo o no; y así sucesivamente.

Los resultados que obtenemos en las clases son muy positivos. Los estudiantes pueden elegir cuáles son las opciones (no solo dos, sino más de dos, como por ejemplo en el lanzamiento de un dado) y comprueban con ejemplos más cercanos cómo la probabilidad “converge” a la probabilidad teórica, y esto después se traduce en que realizan los ejercicios con menos dudas.

Es importante destacar que la metodología propuesta no es exclusiva para la enseñanza de la Estadística, sino que se puede aplicar a cualquier asignatura en la que se puedan utilizar datos, funciones matemáticas, gráficos o simulación, y en las que la interacción del estudiante provoque distintos resultados que le ayuden a interiorizar los conceptos.