# Visualizar la variabilidad: una propuesta didáctica para el desarrollo del pensamiento estadístico con ANOVA

**Elena Vázquez¹, Rosa Alcover², Vicente Chirivella²**

¹Centro de Gestión de la Calidad y del Cambio, Universitat Politècnica de València, España  
2Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València, España

## Author keywords

Variabilidad; Estadística; Pensamiento estadístico; Aprendizaje; Educación superior; ANOVA; Visualización; Herramientas interactivas; Lenguaje R

## Abstract

La variabilidad es un concepto central en Estadística y una piedra angular para el desarrollo del pensamiento estadístico. Sin embargo, su comprensión resulta especialmente difícil para muchos estudiantes, que tienden a centrarse únicamente en medidas de centralización (media, mediana, etc). Esta dificultad se debe a múltiples factores tales como la tradición determinista de la enseñanza matemática, la naturaleza abstracta del concepto de variabilidad, y la falta de experiencias directas y adecuadas con datos reales y procesos aleatorios.  
  
En este trabajo se analiza la importancia de la variabilidad en la Estadística, se revisan algunas de las razones didácticas que explican su difícil asimilación en el aula y se propone una aproximación basada en herramientas visuales e interactivas construidas con R, orientadas a hacer visible la existencia y comportamiento de la variabilidad en los datos y facilitar su análisis.  
  
Como caso de aplicación, se plantea una propuesta didáctica centrada en el Análisis de la Varianza (ANOVA), utilizando visualizaciones dinámicas para ilustrar la descomposición de la variabilidad total en variabilidad explicada por el factor y variabilidad residual. Esta propuesta se contextualiza en la asignatura Modelos Estadísticos para la Toma de Decisiones II del Grado en Ciencia de Datos de la Universitat Politècnica de València, y tiene como objetivo reforzar la comprensión conceptual del ANOVA, promover el razonamiento estadístico y facilitar la toma de decisiones basada en datos.

## 1. Introducción

La variabilidad es el eje estructurante del razonamiento estadístico y del modelo lineal general. La literatura en educación estadística subraya que el alumnado suele privilegiar la media y desatender la variación, lo que dificulta la interpretación de inferencias y contrastes (Cooper & Shore, 2008; Castro Sotos et al., 2007). El Análisis de la Varianza (ANOVA) ofrece un marco privilegiado para abordar esta cuestión, al descomponer la variabilidad total en componentes explicada y residual, conectando directamente con la lógica del test F. Para favorecer un aprendizaje significativo, proponemos ANOVAlab, una herramienta interactiva en R/Shiny orientada a 'hacer visible' la variabilidad y a promover el pensamiento estadístico en contextos de educación superior.

## 2. Metodología

### 2.1. Revisión teórica y fundamentos metodológicos

La propuesta se apoya en el modelo lineal general y en la descomposición de la Suma de Cuadrados Total (SCT) en Suma de Cuadrados de los factores (SCF) y Suma de Cuadrados Residual (SCR). A nivel didáctico, nos apoyamos en investigaciones sobre concepciones erróneas y dificultades en inferencia y variabilidad (Castro Sotos et al., 2007; Cooper & Shore, 2008) y en enfoques que integran visualización e interacción para mejorar la comprensión (Forbes et al., 2014; Loy, 2020; Wang, Rush & Horton, 2017). Asimismo, atendemos a guías y actividades que proponen usar R para simulaciones y docencia de inferencia (Zhang & Maas, 2019) e integrar R con materiales interactivos que reduzcan la carga cognitiva de la programación (Tucker et al., 2022).

### 2.2. Diseño y estructura de ANOVAlab

ANOVAlab es una aplicación R/Shiny con arquitectura modular:  
• Generador de datos (gen\_data.R): simula diseños balanceados variando medias, varianzas y tamaños muestrales, mostrando la evolución de SCT, SCF y SCR.  
• ANOVA de un factor (mod\_one\_factor.R): visualiza la descomposición geométrica, el estadístico F y los residuos, e incluye intervalos LSD.  
• ANOVA de dos factores (mod\_two\_factors.R): explora efectos simples e interacciones, con análisis de homocedasticidad a partir de residuos².  
  
El diseño sigue buenas prácticas de R (módulos en ficheros R independientes) e integra gráficos con ggplot2/plotly para retroalimentación inmediata. El enfoque visual es coherente con propuestas para introducir figuras conceptuales de ANOVA (Kim, 2017) y con usos de datos motivadores y dashboards docentes (Sisso, Bass & Williams, 2023).

### 2.3. Contexto educativo e implementación

La herramienta se implementa en la asignatura Modelos Estadísticos para la Toma de Decisiones II (Grado en Ciencia de Datos, UPV) en tres etapas: exploración conceptual con simulaciones, aplicación a casos reales de uno y dos factores y reflexión final sobre supuestos y toma de decisiones. La aproximación fomenta el aprendizaje activo y la argumentación basada en datos, alineada con prácticas contemporáneas de visualización e inferencia visual en el aula (Loy, 2020; Forbes et al., 2014).

### 2.4. Evaluación prevista del aprendizaje

Se plantea un diseño mixto pretest–postest (aprendizaje conceptual) y análisis cualitativo de razonamientos (respuestas abiertas, microentrevistas). Hipótesis: ANOVAlab aumentará la comprensión de la descomposición de la variabilidad y la interpretación de interacciones, y mejorará la actitud hacia R cuando se integra de forma progresiva en tareas guiadas (Tucker et al., 2022).

## 3. Resultados esperados y aportaciones

Se esperan mejoras en la retención de la lógica de la descomposición de la variabilidad, en la lectura de tablas ANOVA y en la detección e interpretación de interacciones. Pedagógicamente, ANOVAlab contribuye a superar el enfoque mecanicista y a articular una experiencia de 'ver para comprender' que conecta datos, modelo y contraste.

## 4. Conclusiones

ANOVAlab integra visualización, simulación e interacción para desarrollar el pensamiento estadístico en educación superior. Su diseño modular en R/Shiny y su foco en la variabilidad lo convierten en un recurso replicable y extensible, coherente con los materiales docentes de ANOVA utilizados en la UPV.

## 5. Limitaciones y líneas futuras

Limitaciones: enfoque actual en diseños balanceados y ANOVA clásico; falta de resultados empíricos longitudinales. Futuro: módulos de ANCOVA y modelos mixtos; análisis de efectos no balanceados; informes reproducibles en R Markdown; evaluación adaptativa y despliegue multilingüe.

## Referencias

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students’ misconceptions of statistical inference: A review of the empirical evidence. Educational Research Review. https://doi.org/10.1016/j.edurev.2007.04.001

Cooper, L. L., & Shore, F. S. (2008). Students' misconceptions in interpreting center and variability of data via histograms and stem-and-leaf plots. Journal of Statistics Education. https://doi.org/10.1080/10691898.2008.11889559

Forbes, S., Chapman, J. A., Harraway, J., Stirling, D., & Wild, C. (2014). Use of data visualisation in the teaching of statistics: A New Zealand perspective.

Kim, T. K. (2017). Understanding one-way ANOVA using conceptual figures. Korean Journal of Anesthesiology. https://doi.org/10.4097/kjae.2017.70.1.22

Loy, A. (2020). Bringing visual inference to the classroom. https://doi.org/10.1080/26939169.2021.1920866

Sisso, D., Bass, N., & Williams, I. (2023). Teaching One-Way ANOVA with engaging NBA data and R Shiny within a flexdashboard. Teaching Statistics. https://doi.org/10.1111/test.12332

Tucker, M. C., Shaw, S. T., Son, J. Y., & Stigler, J. W. (2022). Teaching Statistics and Data Analysis with R. Journal of Statistics and Data Science Education. https://doi.org/10.1080/26939169.2022.2089410

Wang, X., Rush, C., & Horton, N. J. (2017). Data Visualization on Day One: Bringing big ideas into intro stats early and often. arXiv preprint.

Zhang, X., & Maas, Z. (2019). Using R as a simulation tool in teaching introductory statistics. Journal on Mathematics Education. https://doi.org/10.29333/iejme/5773

How to cite: Vázquez, E.; Alcover, R.; Chirivella, V. 2025. Visualizar la variabilidad: una propuesta didáctica para el desarrollo del pensamiento estadístico con ANOVA. In: IV Congreso & XV Jornadas de Usuarios de R. Universitat Politècnica de València, Valencia, 2025. https://doi.org/10.4995/CongresoR2025.2025.\* [To be completed by the publisher]