

AovBY: Un Paquete de R para aplicación y visualización de ANOVA bayesiano

by Johny J. Pambabay-Calero, Mauricio J. Rojas-Campuzano, Sergio A. Bauz-Olvera, and Omar H. Ruíz-Barzola

Abstract El análisis de varianza es una técnica estadística muy usada en diseño de experimentos y distintas áreas de investigación, por lo general se ha modelizado mediante un enfoque clásico o frecuentista. Con el poder computacional que se cuenta en la actualidad, el enfoque bayesiano resulta una importante herramienta estadística en lo relacionado al contraste de hipótesis. No obstante, el conformismo con las técnicas clásicas, el desconocimiento de la estadística bayesiana y la falta de software de fácil uso son un obstáculo para su frecuente aplicación. En este trabajo se propone un paquete estadístico reproducible en R, el cual presenta opciones para realizar un análisis de varianza de manera clásica (frecuentista), y bayesiana cuando los supuestos del enfoque frecuentista no se cumplen, o cuando se requiere un nivel de inferencia más específico como cuantificar la evidencia que proporciona un conjunto de datos para una hipótesis determinada, con la posibilidad de aportar a la comprensión del rechazo o no de las hipótesis estadísticas planteadas, a través de gráficos interactivos presentados en un panel de Shiny emergente.

Introducción

En el contexto de aplicación de métodos estadísticos, se han valorado muchas técnicas, cada una con diversas especificaciones, supuestos, ventajas y desventajas, según el marco de datos a la que es sometida. El caso del Análisis de Varianza (ANOVA) es especial, ya que se ha mantenido como un modelo fenomenalmente popular (Kozak and Piepho, 2018). Es importante recordar la variabilidad que existe entre los perfiles de los usuarios del modelo, ya que, así como puede ser usado por investigadores o estadísticos, son muchos los casos en que personas, que no acostumbran el estudio de la teoría estadística, lo usan. Esta es una potencial razón para que ANOVA sea objeto de múltiples violaciones en sus supuestos. Generalmente se asume que los residuos se distribuyen normalmente con media 0 y varianza σ^2 para estimar los parámetros desconocidos en el modelo (Çelik and Şenoğlu, 2018). Probablemente, este sea el supuesto más conocido, aún así, su validación suele manejarse incorrectamente. Es común ver pruebas de normalidad aplicadas a los datos brutos y no a los residuos (Kozak and Piepho, 2018). De igual manera existen casos en que no se evalúa de la mejor manera el supuesto de la homocedasticidad. Otros escenarios que se deben considerar son la presencia de valores potencialmente atípicos (Barnett and Lewis, 1994) y los valores perdidos (Rubin, 1987). Una vez que se demuestra que la prueba es aplicable al marco de datos, por medio de la validación de los supuestos, se pueden realizar inferencias confiables.

Si los supuestos no se cumplen, en especial el supuesto de normalidad, el camino inmediato que los investigadores suelen tomar es elegir una de las pruebas no paramétricas (Flores et al., 2017). En el caso de ANOVA, para los mismos fines, se utiliza Kruskal-Wallis, esta no asume la normalidad en los residuos (Ostertagova et al., 2014). Sin embargo, estas pruebas no paramétricas, aunque se dice que no están ligadas a condiciones específicas, tienen un supuesto fundamental, la simetría, que muchas veces no es considerado (Duller and Vorhauer, 2020). Si la media, la mediana y la moda tienen el mismo valor, se puede afirmar que la distribución estadística de los datos es simétrica, lo que agiliza las inferencias estadísticas y aplicar pruebas con medianas y rangos, se vuelve una opción plausible. Si esto no se cumple, es una buena práctica buscar otros métodos que provean inferencias más confiables. Adicional a esto, se sabe que las pruebas no paramétricas, al pertenecer a la estadística clásica, tienen como potencial punto de conclusión al valor-p, y que el mismo, al aumentar el tamaño muestral, puede sugerir significación estadística por alguna diferencia, aunque esta sea pequeña (Jiménez-Paneque, 2016).

Las dificultades presentadas por la estadística frecuentista pueden ser resueltas, cada una con un método diferente, pero es posible abordar de manera integrada todas estas dificultades, con el enfoque bayesiano del análisis de varianza (Wedel and Dong, 2020).

La inferencia bayesiana provee más oportunidades a los usuarios, permitiendo que el conocimiento se actualice en función de los datos y el conocimiento previo que se tiene sobre un caso específico. Incluso puede monitorear la evidencia hasta que el resultado sea lo suficientemente convincente o los recursos disponibles se hayan agotado (Wagenmakers et al., 2018). Evidentemente utilizar técnicas de inferencia bayesiana no es una práctica igual de frecuente que el uso de inferencia clásica, esto puede deberse a tres causas:


- Conformidad con los métodos clásicos.

- Desconocimiento de los modelos bayesianos.
- Ausencia de complemento computacional.

Es claro que estos tres puntos no son excluyentes, debido a que la probabilidad de que alguien se sienta conforme con los métodos clásicos puede aumentar si se le añade el desconocimiento de los métodos de enfoque bayesiano o si no se tiene un software que canalice el conocimiento teórico y permita llevarlo a la práctica con relativa facilidad. Se conoce que el análisis de varianza bayesiano aborda de mejor manera las dificultades ocasionadas por la falta de validez en los supuestos de la estadística clásica, en este punto ya es considerable descartar la primera causa presentada anteriormente, a esto se le suman más ventajas. Mientras que en ANOVA clásico se concluye considerando el valor-p y el nivel de significancia determinado, teniendo como opciones *Se rechaza la hipótesis nula* o *No existe evidencia para rechazar la hipótesis nula*, el enfoque bayesiano permite cuantificar cuánta evidencia proporciona un conjunto de datos para una hipótesis (Wedel and Dong, 2020). El factor de Bayes, cuantifica el rendimiento predictivo relativo de los modelos rivales, además no hay necesidad de acotar el valor del factor de Bayes. A diferencia del valor-p, con el factor de Bayes no hay necesidad de establecer un punto crítico o zona de rechazo, para realizar un contraste de hipótesis (van den Bergh et al., 2020).

Un impedimento importante para la adopción generalizada de las pruebas bayesianas es posiblemente la falta de software fácil de usar para los problemas estadísticos corrientes (Wagenmakers et al., 2018). Actualmente, muchos investigadores o personas que buscan la aplicación del análisis de varianza, con suficiente conocimiento teórico o no, tienen algunas herramientas a disposición, la función *aov* del paquete **stats** en R (Collyer and Adams, 2018), el módulo de Análisis de varianza de STATA (StataCorp., 2019), el procedimiento ANOVA en SAS (SAS Institute Inc., 1999), entre otras. Para este artículo se usará el software estadístico R (R Core Team, 2012). La idea de adquirir otros programas estadísticos que tengan implementos de inferencia bayesiana, como JASP (van den Bergh et al., 2020), no es desatinada, pero, ¿Qué pasaría si se presenta un complemento estadístico interactivo que pueda reproducirse desde R, para la prueba ANOVA Bayesiana? La probabilidad de que la persona que está leyendo este artículo se sienta totalmente conforme con los métodos clásicos quizás se reduce. Hasta hace poco, estas pruebas no se habían implementado en ningún software, y mucho menos en software fácil de usar (Wagenmakers et al., 2018). La pregunta deja de ubicarse en un sentido utópico, ya que ese complemento ahora existe, la interactividad que posee lo hace amigable y los procedimientos bayesianos lo hacen confiable.

A continuación, se enuncia el modelo ANOVA clásico y el modelo ANOVA bayesiano. Para terminar se presenta un ejemplo con la base de datos de un experimento donde se analizó la capacidad de la planta vetiver (*Chrysopogon zizanioides*) para remover del agua cinco productos farmacéuticos catalogados como contaminantes (Checa-Artos et al., 2021), para este ejemplo solo se considerará a la sustancia contaminante SULFAMETOXAZOLE. El diagnóstico de las pruebas estadísticas de estos datos se presenta en la Tabla 1. Para la decisión, se considera un α (Error tipo 1) = 0.05.



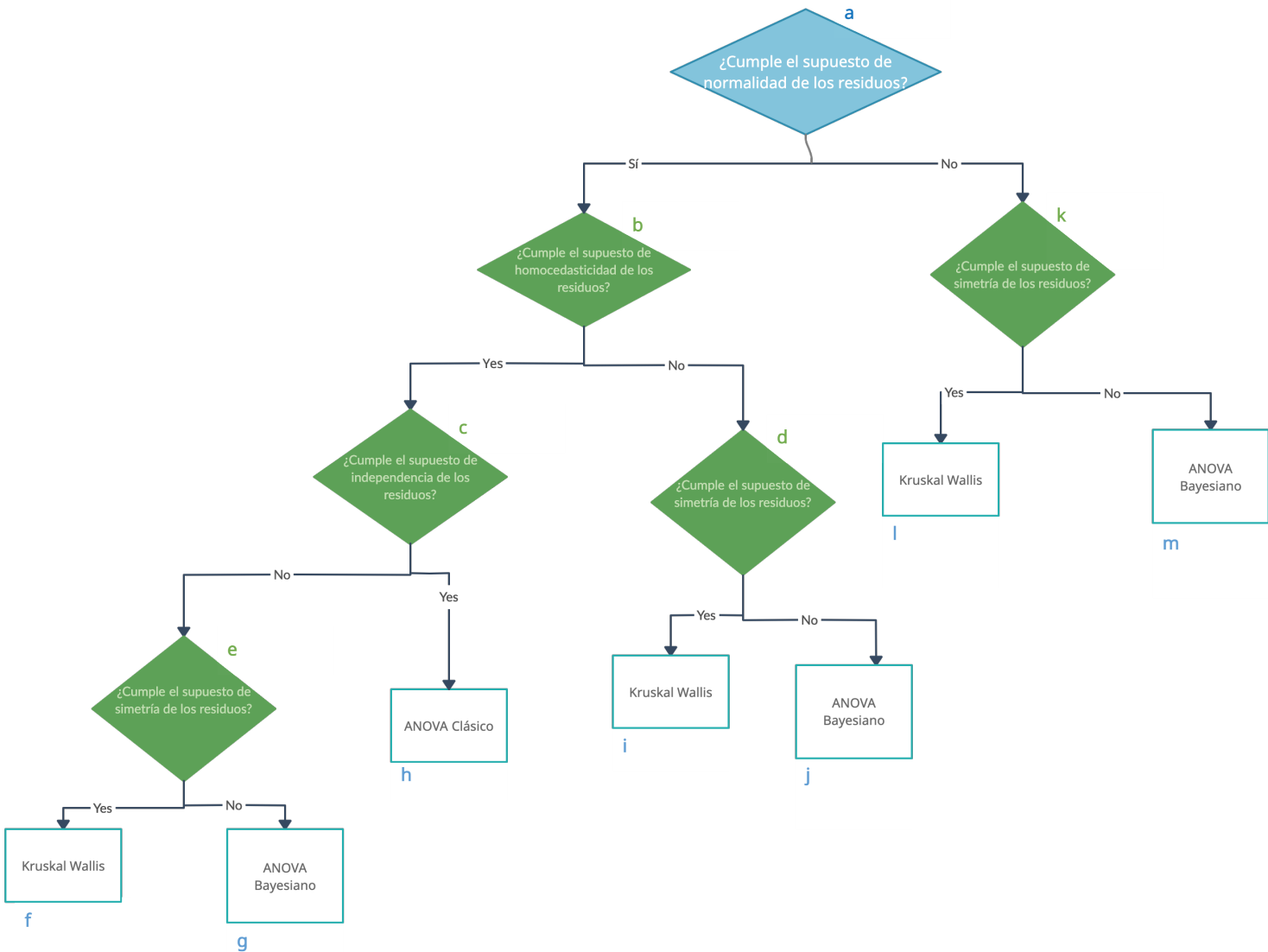
Supuestos	Prueba	Valor-p	Valor	Decisión
Normalidad de los residuos	Kolmogorov-Smirnov	0.3832		Sí cumple
	Shapiro-Wilk	0.3865		
Homocedasticidad de los residuos	Bartlett Test	$< 2.2 \times 10^{-16}$		No cumple
Simetría de los residuos			0.1290 *	No cumple
Independencia de los residuos	Durbin Watson	0.006		No cumple

* Coeficiente de asimetría

Table 1: Evaluación de supuestos

Para poder aplicar ANOVA clásico, es necesario que estos supuestos de la Tabla 1 se cumplan, sin embargo, no es el caso, ya que el único supuesto que cumple es la normalidad, mientras que los supuestos de homocedasticidad, simetría e independencia no se cumplen. Para decidir qué camino tomar, ante el inconveniente de incumplimiento de supuestos, se diseñó un diagrama de decisión, que facilita la selección de la técnica estadística de forma sistemática, en función del cumplimiento o incumplimiento de supuestos.

Figure 1: Diagrama de flujo: Supuestos y técnicas estadísticas.



La Figura 1 muestra un diagrama de flujo, donde se mencionan los supuestos y se establecen las vías disponibles según su cumplimiento, terminando en la técnica aplicable al marco de datos. El recorrido comienza en el rombo de decisión **a**, donde se cuestiona el supuesto de normalidad de los residuos, si este se cumple, se procede al rombo **b** donde se pregunta si el supuesto de homocedasticidad de los residuos se cumple, si este supuesto es validado, se sigue al rombo **c**, donde se cuestiona el supuesto de independencia de los residuos. Si el supuesto es favorable es una buena práctica aplicar el modelo ANOVA clásico. Sin embargo, si el supuesto de independencia, homocedasticidad o normalidad, mencionados en los rombos de decisión **c**, **b** y **a** respectivamente, no se cumplen, se procede a evaluar el supuesto de simetría de los residuos (rombos **d**, **e** y **k**), si este se cumple, la técnica recomendada es Kruskal Wallis, si no se cumple, es recomendable realizar el análisis de varianza desde el enfoque bayesiano. Es importante mencionar que si se desea evitar la evaluación de supuestos, se puede empezar directamente con el ANOVA bayesiano.

Usando la Tabla 1 y el diagrama de flujo, revisado en la Figura 1, se tomará la decisión de que técnica es la más apropiada, sometiendo el conjunto de datos a los supuestos sugeridos en el diagrama, de forma sistemática. Se comienza evaluando el supuesto de normalidad de los residuos, el diagnóstico del marco de datos, observado en la Tabla 1, indica que sí se cumple, esto da paso a la siguiente incógnita, que trata sobre la homocedasticidad, la Tabla 1 indica que no se cumple este supuesto. Desde esta ubicación, se descarta la aplicación de ANOVA clásico. Siguiendo con la ruta que sugiere el diagrama, se pregunta sobre el cumplimiento del supuesto de simetría, que tampoco se cumple, por lo tanto no se debería considerar el equivalente no paramétrico Kruskal Wallis. El camino sugerido por el diagrama, es la aplicación del análisis de varianza con enfoque bayesiano.

Materiales y Métodos

ANOVA de una vía

Cuando solo tenemos un factor, el modelo ANOVA puede ser denotado de la siguiente manera:

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad ; \quad i = 1, \dots, k, \quad j = 1, \dots, n_i \quad (1)$$

donde

i representa el i -ésimo tratamiento del factor y el número de tratamientos total es denotado por k .

j representa la observación, el número máximo de j es n_i , que representa el número de observaciones total en el tratamiento i .

Y_{ij} es el valor que toma la variable de respuesta en la j -ésima observación del i -ésimo tratamiento.

μ_i es la media del i -ésimo tratamiento.

ϵ_{ij} es el error aleatorio, independiente e idénticamente distribuido de forma normal con media 0 y varianza σ^2 .

El análisis de varianza clásico se enfoca en la prueba de las siguientes hipótesis:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu \quad \text{versus} \quad H_1 : \mu_i \neq \mu$$

La hipótesis alterna indica que al menos una μ_i (Media del tratamiento i) es diferente a las demás.

El modelo de análisis de varianza, compara las variaciones entre grupos (tratamientos) y dentro de los grupos (Ostertagova et al., 2013), se sabe que si no existe efecto del factor en la variable de respuesta, estas variaciones no serán diferentes, caso contrario, la diferencia entre estas se reflejará en el estadístico F y se dará paso al rechazo de la hipótesis nula.

Para poder obtener el estadístico F , es necesario obtener las fuentes de variación y sus grados de libertad, la suma cuadrada de tratamientos $SSTR$ con $k-1$ grados de libertad, la suma cuadrada de los residuos SSE con $n-k$ grados de libertad y la suma cuadrada total $SSTO$ con $n-1$ grados de libertad.

El estadístico F es el siguiente:

$$F = \frac{SSTR/(k-1)}{SSE/(n-k)} \quad (2)$$

Se sabe que bajo la hipótesis nula H_0 el estadístico de prueba F sigue una distribución $F_{k-1, n-k}$ y bajo la hipótesis alterna se distribuye como una F no central $F_{k-1, n-k, \lambda}$, donde λ es el parámetro de desentralización (Solari et al., 2008).

ANOVA Bayesiano

Se conoce que el objetivo principal de la inferencia bayesiana es la actualización del conocimiento mediante las observaciones, a esto se le puede llamar *aprendizaje* (van den Bergh et al., 2020). En el análisis de varianza de una vía se plantea inicialmente la hipótesis nula, que indica que no hay efecto del factor en la media de la distribución de la variable de respuesta, y la hipótesis alternativa, que contempla los escenarios en que esto no sucede. Con ANOVA bayesiano se puede determinar que hipótesis es más probable, comparando las probabilidades posteriores de cada modelo candidato. Si se acota el alcance a modelos de una vía, los modelos candidatos son el modelo nulo Θ_0 , que considera una sola media constante, y el modelo que considera el efecto de los tratamientos en la variable dependiente Θ_1 . Para llegar a la comparación de modelos posteriores, se debe partir del conocimiento previo e incluir la verosimilitud o los datos.

Priori

Una vez que se identifican las hipótesis se determina su probabilidad a priori o previa $\pi(\Theta_i)$. Al tener dos hipótesis, usualmente se calcula la razón entre ellas:

$$\frac{\pi(\Theta_1)}{\pi(\Theta_0)} \quad (3)$$

tal que $\pi(\Theta_1)$ y $\pi(\Theta_0)$ son las probabilidades a priori de las hipótesis alternativa y nula, respectivamente.

Al tener esta razón, se tienen 3 escenarios (RENDON-MACIAS et al., 2018).

1. La razón es igual a 1, esto implica que las dos hipótesis son igual de probables.
2. La razón es mayor a 1, se tiene información previa que indica que la hipótesis alternativa es más probable.
3. La razón es menor a 1, se conoce previamente que la hipótesis nula es más probable.

Verosimilitud

Una vez que se tienen las probabilidades previas, para poder actualizar el conocimiento, se requiere la información que proveen los datos (y), la distribución muestral o distribución de los datos $p(y|\Theta_i)$. La inferencia bayesiana indica que dos modelos de probabilidad $p(y|\Theta)$ que tengan la misma función de verosimilitud producen la misma inferencia para cualquier hipótesis Θ (Gelman et al., 2013).

Posterior

Si condicionamos el valor conocido de los datos, utilizando la teoría de probabilidad condicional o regla de Bayes, se puede obtener la densidad posterior $p(\Theta|y)$ (Gelman et al., 2013).

$$p(\Theta|y) \propto \pi(\Theta)p(y|\Theta) \quad (4)$$

donde,

$\pi(\Theta)$ es la probabilidad a priori y

$p(y|\Theta)$ la verosimilitud.

Factor de Bayes

El Factor de Bayes (FB), también conocido como razón de verosimilitud, utiliza el conocimiento previo para transformarlos en inferencias, es decir, determinar que una conclusión sea correcta o incorrecta en función de probabilidades (Goodman, 1999). Al ser una razón, puede expresarse de la siguiente manera:

$$FB_{01} = \frac{p(y|\Theta_0)}{p(y|\Theta_1)} \quad (5)$$

Donde $p(y|\Theta_0)$ y $p(y|\Theta_1)$ son las distribuciones de los datos bajo la hipótesis nula y alterna, respectivamente. Si se requiere FB_{10} , basta con obtener el recíproco de FB_{01} , $FB_{01} = \frac{1}{FB_{10}}$ (Rouder et al., 2012). Se puede utilizar esta razón de verosimilitud, para encontrar una razón que involucre el conocimiento posterior, dándole una mirada a la ecuación (4).

$$\frac{p(\Theta_0|y)}{p(\Theta_1|y)} = \frac{p(y|\Theta_0) \pi(\Theta_0)}{p(y|\Theta_1) \pi(\Theta_1)} = FB_{01} \frac{\pi(\Theta_0)}{\pi(\Theta_1)} \quad (6)$$

donde $\frac{\pi(\Theta_0)}{\pi(\Theta_1)}$ es la razón planteada en la ecuación (3).

De esta manera, se puede concluir con el conocimiento posterior. El factor de bayes es el responsable del cambio en las creencias, reflejadas en la razón de las probabilidades posteriores (Rouder et al., 2012).

Ejemplo ilustrativo

Se parte de un estudio en el que, con 5 productos farmacéuticos, se prepararon soluciones acuosas usando metanol ultrapuro a cuatro concentraciones: 3 mg/L, 6 mg/L, 9 mg/L, 12 mg/L y 4 tiempos de muestreo. Posteriormente se introdujo la especie vetiver (*Chrysopogon zizanioides*), bajo condiciones específicas el experimento se llevó a cabo. La variable de respuesta es el porcentaje de remoción de estos contaminantes. Si el lector desea profundizar en esta investigación, realizada con superficies de respuesta, puede ver también Checa-Artos et al. (2021).

En este ejemplo, se utiliza la sustancia contaminante *SULFAMETOXAZOLE*. Se delimita como factor a la concentración del contaminante (en ppm).

Tratamientos	Unidades Experimentales
Concentración 3	4
Concentración 6	2
Concentración 9	2
Concentración 12	4

Table 2: Descripción de los datos del ejemplo ilustrativo

La Tabla 2 evidencia que no se trata de un diseño balanceado, ya que el número de unidades experimentales no es el mismo en cada nivel del factor. Para analizar estos datos se usará el paquete **AovBY**, que será descrito a continuación.

El paquete

El paquete **AovBY** incluye la función `aovbay()`, donde se tiene un único argumento, **data**, que corresponde al conjunto de datos. La función puede ser llamada de la siguiente manera:

```
library(AovBY)
data(PollutantRemoval)
aovbay(PollutantRemoval)
```

La salida de esta sentencia es un panel de shiny, donde se muestra una ventana dedicada a la carga o actualización del conjunto de datos. Esto se puede observar en la Figura 2.

Incluso, es posible acceder al panel de shiny, llamando a la función sin especificar ningún argumento, de la siguiente forma:

```
library(AovBY)
aovbay()
```

El usuario puede ingresar la base de datos desde el panel, sin necesidad de haberla ingresado en R. La salida puede verse en la Figura 3.

Figure 2: Ventana de carga de datos, carga desde la función.

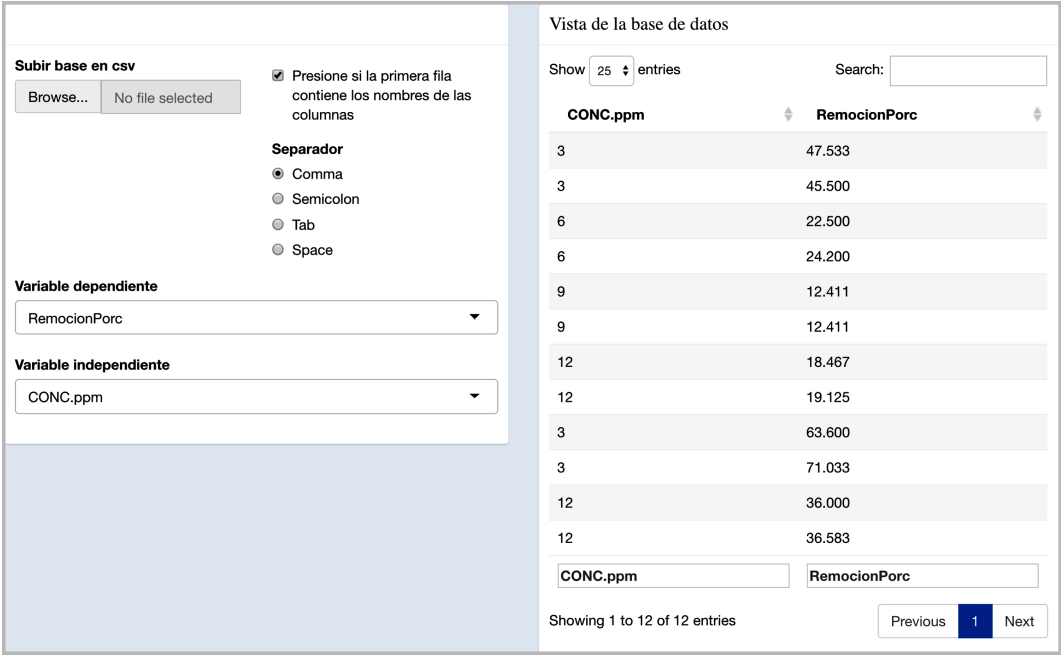
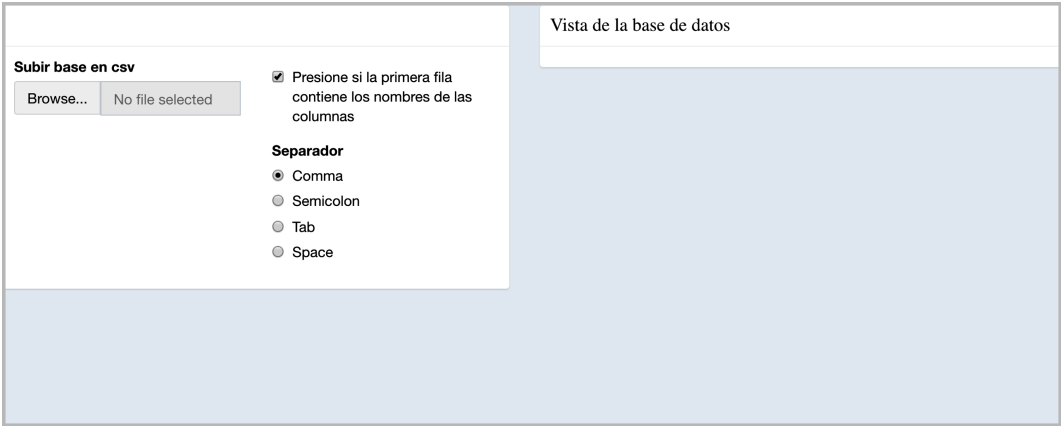


Figure 3: Ventana de carga de datos, carga desde el panel.



La función no contiene más argumentos, no obstante, esto no significa que el usuario recibe una salida fija, inmodificable, ya que se tiene la posibilidad de cambiar argumentos en las pruebas estadísticas, en las simulaciones, entre otros, dentro de la aplicación emergente, de forma amigable y reactiva, esto último implica que si se cambia algún argumento, la salida se renovará y se presentará en función del argumento nuevo, sin necesidad de volver a correr la función.

Se presenta un menú, donde se pone a disposición del usuario los supuestos, una ventana con la salida del ANOVA clásico, otra con Kruskal Wallis y otra con ANOVA bayesiano, visible en la Figura 4.

Figure 4: Menú del panel shiny emergente.



Resultados

Para la descripción de los resultados se utiliza el diagrama de técnicas según el cumplimiento de supuestos, revisado en la Figura 1.

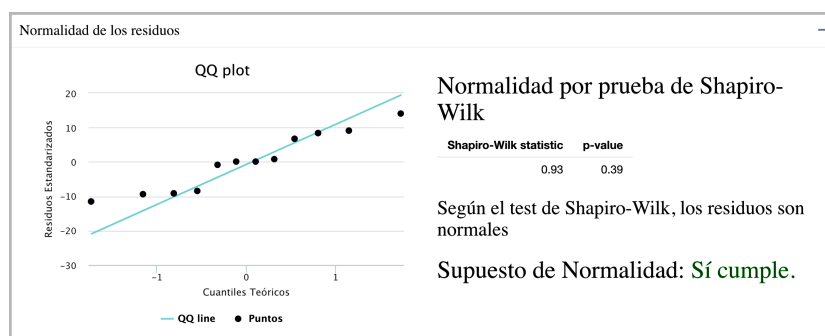
Supuestos

El paquete permite evaluar los supuestos, de forma gráfica y mediante pruebas estadísticas. Se presenta la conclusión, basada en el valor-p del test aplicado y el α (error tipo 1) con el que el usuario desee realizar los contrastes, si no se desea cambiar este valor, se mantiene $\alpha = 0.05$. Posteriormente, se hace un recuento de los supuestos validados y se sugiere la técnica apropiada.

Normalidad de los residuos

Para la evaluación del supuesto de normalidad de los residuos, se usa la prueba de Kolmogórov-Smirnov con la corrección Lilliefors, en el caso de que el tamaño muestral sea mayor a 30, y la prueba Shapiro-Wilk, cuando el tamaño de muestra es menor a 30 (Gandica de Roa, 2020). La función *aovbay()* evalúa esta condición y muestra la prueba de normalidad de residuos adecuada. Se presenta la decisión, basada en el valor-p y el α (error tipo 1) seleccionado por el usuario, adicionalmente, se tiene el gráfico Q-Q plot, que aporta a la comprobación de la inferencia.

Figure 5: Normalidad de los residuos, salida del paquete



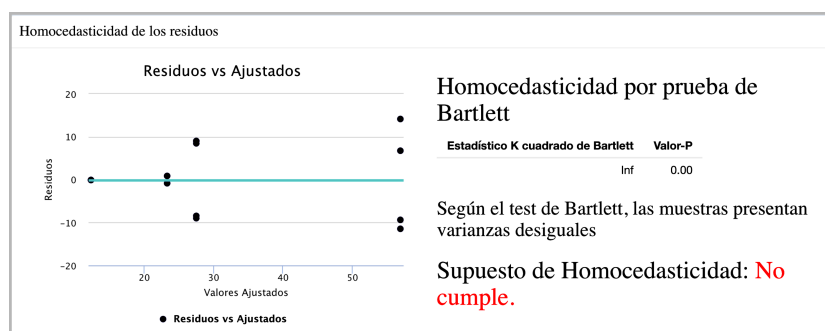
La Figura 5 presenta la salida que proporciona el paquete al evaluar la normalidad de los residuos con los datos del ejemplo ilustrativo. La prueba Shapiro-Wilk sugiere que los residuos, si pueden provenir de una distribución normal. El gráfico Q-Q plot evidencia esta inferencia.

Siguiendo la ruta propuesta en la Figura 1, el siguiente paso es verificar si el supuesto de homocedasticidad de los residuos del modelo ANOVA, mostrado en la ecuación 1, se cumple.

Homocedasticidad de los residuos

El supuesto de homocedasticidad de los residuos es verificado con el test de Bartlett, que tiene como objetivo probar que las varianzas en las k poblaciones, determinadas por los tratamientos, sean las mismas (Aslam, 2020). Se tiene también el gráfico de Residuos versus Valores ajustados, para identificar posibles secciones en las que la varianza no sea la misma.

Figure 6: Homocedasticidad de los residuos, salida del paquete



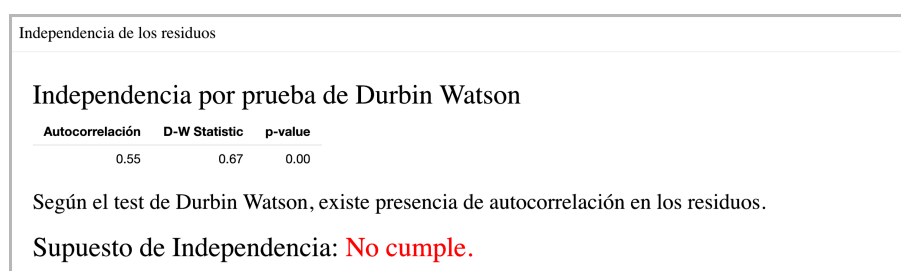
En la Figura 6, se observa el resultado del test **de Bartlett**, siendo el valor-p muy cercano a cero, concluyendo que no se cumple el supuesto de homocedasticidad. Esto es confirmado con el gráfico de residuos versus valores ajustados, donde se aprecia el incremento en la varianza según se avanza hacia la derecha, mostrando heterocedasticidad.

Al no cumplir este supuesto, el diagrama presentado en la 1 sugiere probar el supuesto de la simetría. Sin embargo, para terminar los supuestos relacionados con estadística paramétrica, se presenta el supuesto de independencia de los residuos.

Independencia de los residuos

El supuesto de independencia es verificado con el test de Durbin Watson. Se contrasta la presencia de autocorrelación en los residuos del modelo lineal, esto sucede cuando los residuos secuenciales están correlacionados, de modo que si se grafica los residuos vs el tiempo, se obtienen patrones gráficos. Este comportamiento se lo denomina *autoregresivo de primer orden*.

Figure 7: Independencia de los residuos, salida del paquete

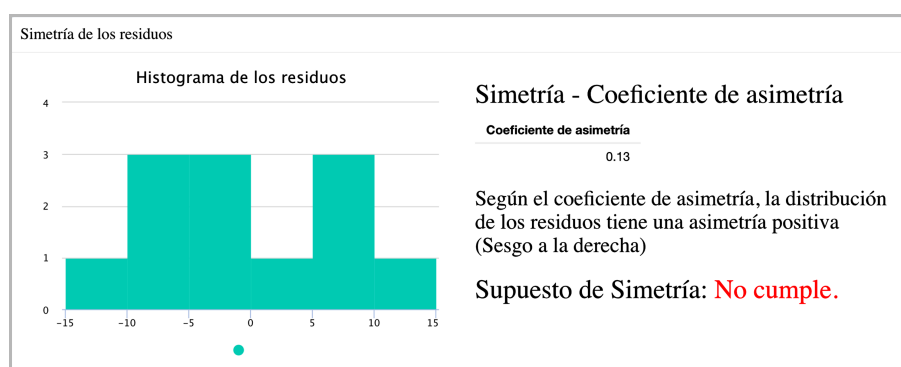


En la Figura 7, se observa la prueba de Durbin Watson, el valor-p es muy cercano a cero, se puede inferir que es muy probable que existan autocorrelaciones, es decir, que no se cumple el supuesto de independencia.

Simetría de los residuos

El supuesto de la simetría es probado con el coeficiente de asimetría. Se sabe que una distribución simétrica tiene coeficiente de asimetría igual a 0, por lo tanto, los casos en los que esto no se cumple son considerados los responsables de la no validación del supuesto. Además se incluye el histograma de los residuos, para verificar si la simetría está presente de forma inmediata.

Figure 8: Simetría de los residuos, salida del paquete



La Figura 8 muestra el coeficiente de asimetría, al ser positivo, se tiene un sesgo en la cola derecha. Por lo tanto el supuesto de simetría no se cumple y por consecuente, se descarta el uso de estadística no paramétrica.

Selección de técnica apropiada

Una vez que se han analizado supuestos, el siguiente paso es seleccionar la técnica apropiada. Con este fin, el paquete provee una sugerencia al usuario.

Figure 9: Técnica apropiada, salida del paquete

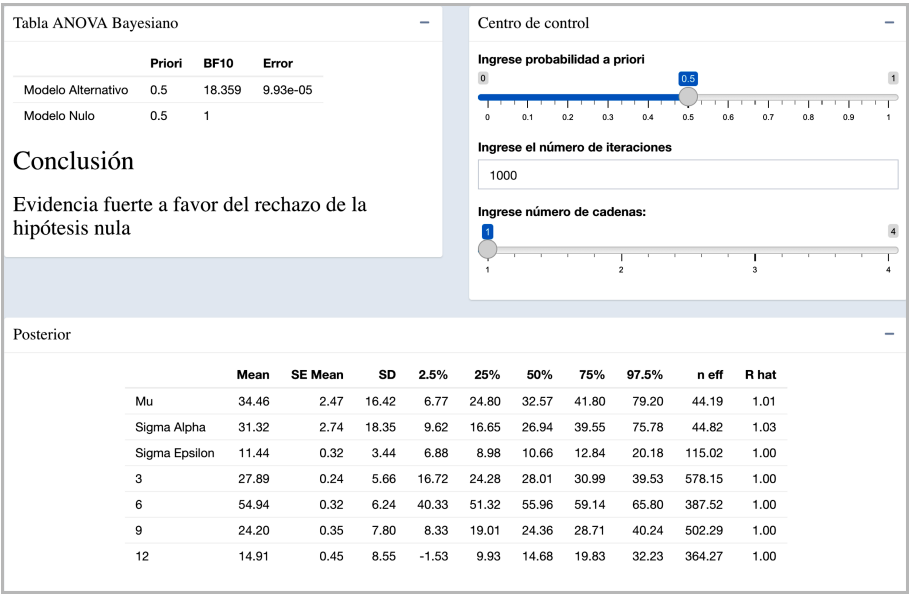


En función del cumplimiento de supuestos, se presentan las técnicas disponibles y la que es considerada la más apropiada. La Figura 9 muestra la forma en la que el paquete sugiere qué técnica es la más apropiada, en función de los supuestos.

Análisis de varianza bayesiano

Para la obtención de los resultados del análisis de varianza con enfoque bayesiano se utilizó los paquetes *BayesFactor* (Morey and Rouder, 2018) y *rstan* (Stan Development Team, 2020). Se parte de la especificación de la priori del modelo alternativo. Se plantea un escenario en el que no se conoce si el factor tiene efecto sobre la variable de respuesta, por lo tanto, la probabilidad a priori, permanecerá con el valor de 0.5 . El número de iteraciones, para las simulaciones Monte Carlo ¹, inicialmente está fijado en 1000, como número mínimo de referencia (Flegal et al., 2008). Sin embargo, si el usuario desea aumentar el valor de este argumento, tiene la posibilidad de hacerlo. Del mismo modo, es posible elegir el número de cadenas Markov ², aumentar este número provee al usuario un efecto de multiplicación del tamaño muestral, disminuyendo el error muestral y la autocorrelación, además esto disminuye la probabilidad de que la cadena se centre en un óptimo local.

Figure 10: Análisis de varianza bayesiano, salida del paquete



La Tabla del análisis de varianza bayesiano, visible en la Figura 10, presenta las probabilidades a priori que fueron consideradas, del modelo alternativo y el modelo nulo, posteriormente presenta el factor de bayes, $BF_{10} = 18.359$, esto quiere decir que el modelo alternativo, que contempla el efecto de la concentración de SULFAMETOXAZOLE en el porcentaje de remoción del contaminante en presencia de la especie vetiver (*Chrysopogon zizanioides*), es 18.359 veces más probable que el modelo nulo, que sólo contempla la gran media, por lo tanto, según la escala de valores de interpretación del factor de bayes, revisada en Gelman et al. (2013), se tiene una evidencia fuerte a favor del rechazo de la hipótesis nula.

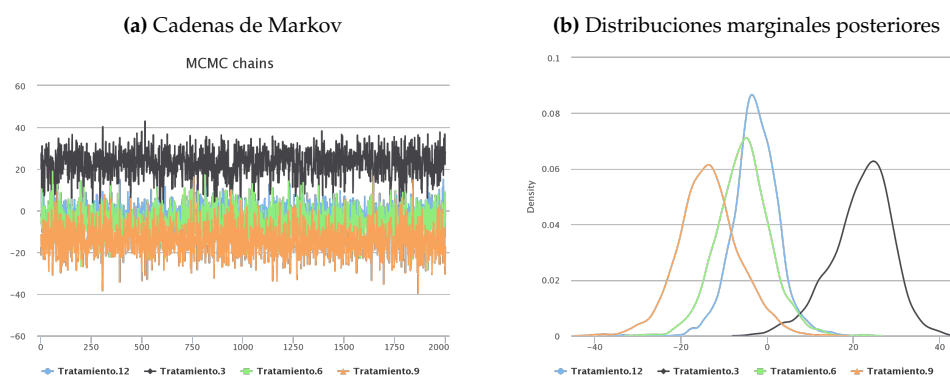
Como parte del método se utiliza el método de cadena de Markov , que mediante simulación

¹La cadena de Markov es una herramienta popular en el entorno de sistemas aleatorios, donde la simulación de cada valor está atado únicamente al valor inmediatamente anterior (Li et al., 2020).
²La cadena de Markov Monte-Carlo (MCMC) es un método muy usado para la estimación de distribuciones posteriores en inferencia bayesiana (Rouder et al., 2018).

Monte Carlo, presentan la distribución posterior de la media, varianza y los tratamientos. Esto puede observarse en la Figura 11.

Adicionalmente, en la Figura 10, se tiene una Tabla donde se muestra la media posterior $Mean$, el error estándar de la media del parámetro $SE\ Mean$, la desviación estándar del parámetro SD , los percentiles 2.5, 25, 50, 75, 97.5, el tamaño de muestra efectivo estimado n_{eff} ³ y el estadístico de reducción de escala potencial $Rhat$ ⁴, para las distribuciones posteriores de la Media general, la desviación entre grupos, desviación dentro de grupos, y cada tratamiento.

Figure 11: Cadenas de Markov y distribuciones posteriores, salida del paquete



En la Figura 11 se puede observar las cadenas de Markov y las distribuciones estadísticas que genera el método MCMC. Se observa que, con el tratamiento que considera una concentración de 3 mg/L, se tiene un comportamiento diferente, observable en las curvas de densidad, se tiene un desplazamiento hacia la derecha, mientras que las distribuciones de los tratamientos con 6 mg/L, 9 mg/L y 12 mg/L se encuentran juntas, incluso se traslapan casi en su totalidad. Esto se puede evidenciar también en las cadenas de Markov, mostradas en la Figura 11, gráfico izquierdo, la línea celeste, correspondiente al tratamiento con 3 mg/L de concentración del contaminante se encuentra más arriba que las otras, correspondientes a los demás tratamientos.

Discusión

En este artículo se hace énfasis en las ventajas del uso del factor de bayes ante el valor p , no obstante, la prueba de supuestos estadísticos, contrastes paramétricos y no paramétricos, al formar parte de la estadística clásica, siguen dependiendo de este valor. Incluso, la decisión que se presenta en el panel de shiny, al evaluar los supuestos, es el resultado de la interpretación de la prueba mediante el valor p . Aún cuando se conoce que el factor de bayes es invariante ante transformaciones lineales y es más estable computacionalmente, debido que tiene algoritmos definidos para el muestreo y obtención de las distribuciones posteriores, cuando se trata de inferencia estadística, el valor p es más utilizado y está presente en la mayoría de contrastes estadísticos, según Rouder et al. (2012), esto puede deberse a que no es usual encontrar fórmulas o herramientas computacionales, basadas en factor de bayes, de fácil uso. El paquete **AovBY** rompe ese paradigma, demostrando que es posible incorporar la inferencia bayesiana, el software estadístico R y la interactividad, responsable del fácil uso de la herramienta.

La Tabla 2, que describe los datos del ejemplo ilustrativo, indica que, además de no ser un diseño balanceado, existen pocas unidades experimentales en cada tratamiento. Resulta complicado describir el comportamiento de una población cuando solo hay dos observaciones. Al realizar el análisis de varianza bayesiano, esto ya no es un inconveniente, ya que, al generarse las cadenas de Markov Monte-Carlo, se tiene la distribución posterior de estas poblaciones, pertenecientes a los tratamientos, con el número de iteraciones que el usuario desee realizar. Resulta más fácil y conveniente inferir sobre una población con 1000 (o más) observaciones que con 2, como es el caso de los tratamientos con concentración 6 mg/L y 9 mg/L (Tabla 2). Además, al comparar las distribuciones posteriores con las curvas de densidad, ya no es necesario hacer múltiples comparaciones de medias, como en el caso del análisis de varianza clásico.

³La cadena de Markov generada tiene aproximadamente la misma precisión que se obtendría de una muestra independiente de tamaño n_{eff} (Ruppert, 2011). Mientras mayor sea este valor, mejor.

⁴Cuando $Rhat$ está lo suficientemente cerca de 1, el diagnóstico declara convergencia (Vats and Knudson, 2020).

Los resultados del análisis de varianza con enfoque bayesiano, obtenidos con la herramienta propuesta, reflejan que existen diferencias significativas en los niveles de concentración del contaminante, si se analizan las curvas de densidad posteriores presentadas en la Figura 11, el tratamiento con 3 mg/L, presenta un desplazamiento hacia la derecha, lejos de las demás curvas, comportamiento que se ve también en las cadenas de Markov, presentadas también en la Figura 11 (gráfico izquierdo), en este caso se ve que la línea que corresponde a las simulaciones de la distribución de la población proveniente del tratamiento con 3 mg/L, se encuentra más arriba que las demás. Con esto se puede inferir que el tratamiento con 3 mg/L es responsable de un mayor porcentaje de remoción del contaminante.

Los resultados obtenidos por parte de [Checa-Artos et al. \(2021\)](#), utilizando superficies de respuesta, indican que para encontrar el porcentaje de remoción óptimo sería necesario reducir las concentraciones, en el caso del contaminante Sulfametoxazole, pero por las condiciones del experimento, la concentración 3 mg/L se contempla como la mejor opción para obtener un mayor porcentaje de remoción del contaminante. La conclusión que se obtiene con los resultados del paquete **AovBY** es la misma que se tiene en el estudio de superficies de respuesta.

Si el usuario desea comparar el resultado obtenido con enfoque bayesiano y el que obtiene con estadística clásica, el panel de shiny emergente incluye también el análisis de varianza clásico, con prueba **post hoc** y el equivalente no paramétrico Kruskal Wallis, visibles en las Figuras 12, 13 y 14, respectivamente.

La herramienta propuesta para decidir que técnica utilizar y aplicar, en el caso del análisis de varianza, proporciona a los investigadores una facilidad en la interpretación de los resultados estadísticos mediante el análisis detallado de los datos, y la interactividad lograda con los paquetes *highcharter* y *shiny*, dentro de RStudio. Aunque inicialmente se recomienda la herramienta más apropiada, en función del cumplimiento de supuestos, en el panel de shiny se presentan los 3 enfoques del análisis de varianza, para que los analistas de datos y usuarios de esta herramienta, puedan valorarlo y tomar desiciones oportunas.

Anexos

Figure 12: Anova clásico, salida del paquete

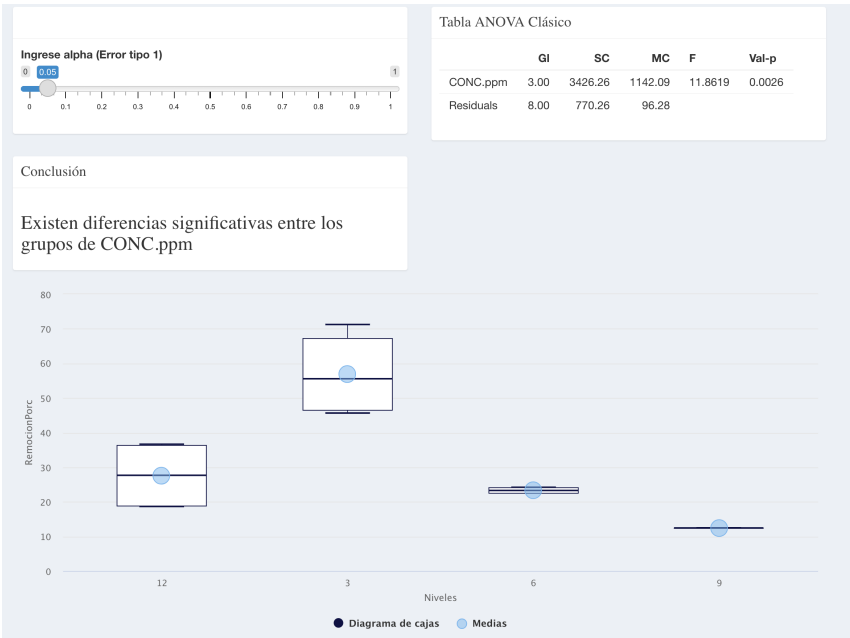


Figure 13: Pruebas Post Hoc (Anova clásico), salida del paquete

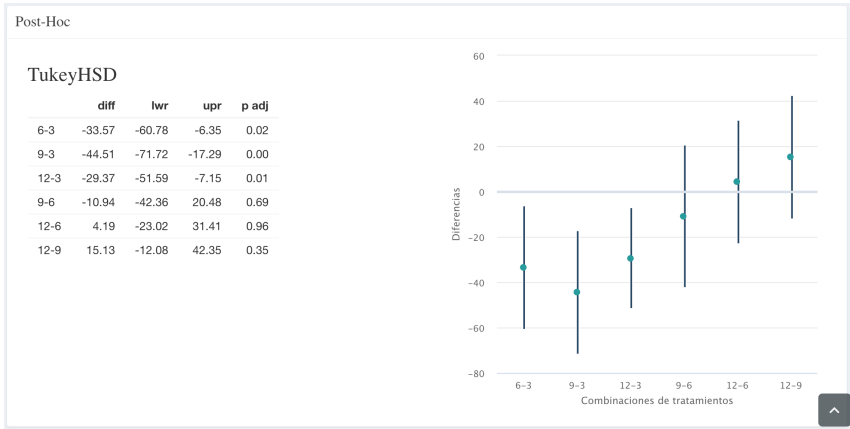
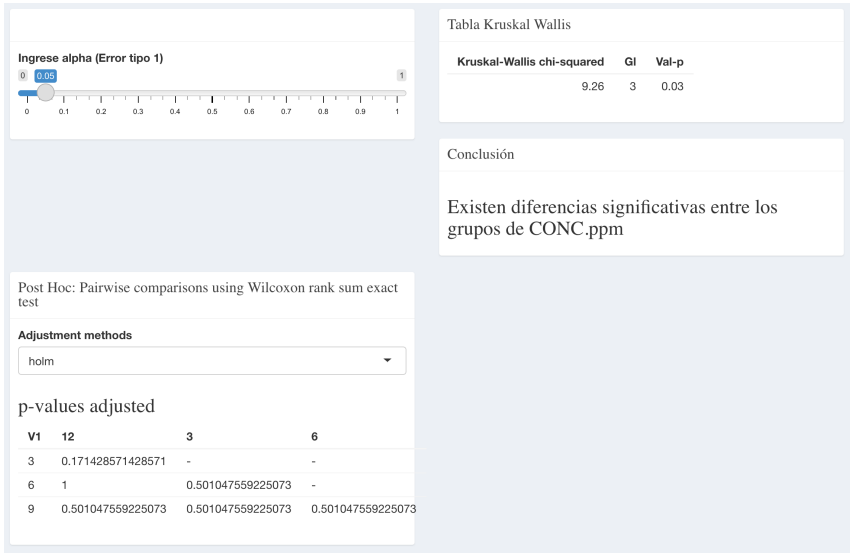


Figure 14: Kruskal Wallis, salida del paquete



Bibliography

M. Aslam. Design of the bartlett and hartley tests for homogeneity of variances under indeterminacy environment. *Journal of Taibah University for Science*, 14(1):6–10, 2020. doi: 10.1080/16583655.2019.1700675. URL <https://doi.org/10.1080/16583655.2019.1700675>. [p8]

V. Barnett and T. Lewis. Outliers in statistical data. (3rd edition.) New York, NY: Wiley, 1994. [p1]

M. Checa-Artos, D. S. del Castillo, V. M. Eulalia, O. Ruiz-Barzola, and M. Barcos-Arias. Remoción de cinco productos farmacéuticos catalogados como contaminantes emergentes en medio acuoso utilizando la especie vetiver (*Chrysopogon zizanioides*). *Bionatura*, 2021. doi: 10.21931/RB/2021.06.01.7. URL <http://dx.doi.org/10.21931/RB/2021.06.01.7>. [p2, 6, 12]

M. L. Collyer and D. C. Adams. Rrpp: An r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution*, 9(7):1772–1779, 2018. doi: 10.1111/2041-210X.13029. URL <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13029>. [p2]

C. Duller and D. Vorhauer. Comparison of four nonparametric tests of symmetry. *AIP Conference Proceedings*, 2293, 2020. doi: 10.1063/5.0026685. URL <https://doi.org/10.1063/5.0026685>. [p1]

- J. M. Flegal, M. Haran, and G. L. Jones. Markov chain monte carlo: Can we trust the third significant figure? *Statistical Science*, 23(2):250–260, 2008. ISSN 08834237. URL <http://www.jstor.org/stable/27645897>. [p10]
- E. Flores, M. G. Miranda, and M. Ángel Villasís. El protocolo de investigación vi: cómo elegir la prueba estadística adecuada. *estadística inferencial*. 64(3):364–370, 2017. doi: 10.29262/ram.v64i3.304. URL <https://doi.org/10.29262/ram.v64i3.304>. [p1]
- E. M. Gandica de Roa. Potencia y robustez en pruebas de normalidad con simulación montecarlo. *Revista Científica*, 5(18):108–119, nov. 2020. doi: 10.29394/Scientific.issn.2542-2987.2020.5.18.5.108-119. URL http://www.indteca.com/ojs/index.php/Revista_Scientific/article/view/468. [p8]
- A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2013. ISBN 9780429113079. URL <http://www.stat.columbia.edu/~gelman/book/>. [p5, 10]
- S. N. Goodman. Toward evidence-based medical statistics. 2: The bayes factor. *Annals of Internal Medicine*, 130(12):1005–1013, 1999. doi: 10.7326/0003-4819-130-12-199906150-00019. URL <https://www.acpjournals.org/doi/abs/10.7326/0003-4819-130-12-199906150-00019>. PMID: 10383350. [p5]
- R. Jiménez-Paneque. El valor de p en entredicho: significación estadística, clínica y práctica. *Medwave*, 2016. doi: 10.5867/medwave.2016.08.6534. URL <https://www.medwave.cl/link.cgi/Medwave/Series/TyC-Estadística/6534.act>. [p1]
- M. Kozak and H.-P. Piepho. What’s normal anyway? residual plots are more telling than significance tests when checking anova assumptions. *Journal of Agronomy and Crop Science*, 204(1):86–98, 2018. doi: <https://doi.org/10.1111/jac.12220>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/jac.12220>. [p1]
- Z. Li, M. Huang, X. Meng, and X. Ge. The limit theorems for function of markov chains in the environment of single infinite markovian systems. *Mathematical Problems in Engineering*, 2020, 2020. doi: 10.1155/2020/8175723. URL <https://doi.org/10.1155/2020/8175723>. [p10]
- R. D. Morey and J. N. Rouder. *BayesFactor: Computation of Bayes Factors for Common Designs*, 2018. URL <https://CRAN.R-project.org/package=BayesFactor>. R package version 0.9.12-4.2. [p10]
- Ostertagova, Eva, Ostertag, and Oskar. Methodology and application of one-way anova. *American Journal of Mechanical Engineering*, 1:256–261, 11 2013. doi: 10.12691/ajme-1-7-21. [p4]
- E. Ostertagova, O. Ostertag, and J. Kováč. Methodology and application of the kruskal-wallis test. *Applied Mechanics and Materials*, 611:115–120, 08 2014. [p1]
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL <http://www.R-project.org/>. ISBN 3-900051-07-0. [p2]
- M. E. RENDON-MACIAS, A. RIOJAS-GARZA, D. CONTRERAS-ESTRADA, and J. D. MARTINEZ-EZQUERRO. Análisis bayesiano. conceptos básicos y prácticos para su interpretación y uso. *Revista alergia México*, 65:285 – 298, 09 2018. ISSN 2448-9190. URL http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-91902018000300205&nrm=iso. [p5]
- J. N. Rouder, R. D. Morey, P. L. Speckman, and J. M. Province. Default bayes factors for anova designs. *Journal of Mathematical Psychology*, 56(5):356–374, 2012. ISSN 0022-2496. doi: <https://doi.org/10.1016/j.jmp.2012.08.001>. URL <https://www.sciencedirect.com/science/article/pii/S0022249612000806>. [p5, 6, 11]
- J. N. Rouder, R. D. Morey, P. L. Speckman, and J. M. Province. A simple introduction to markov chain monte-carlo sampling. *Psychonomic Bulletin & Review*, 25, 2018. doi: 10.3758/s13423-016-1015-8. URL <https://doi.org/10.3758/s13423-016-1015-8>. [p10]
- D. B. Rubin. Multiple imputation for nonresponse in surveys. *Wiley Series in Probability and Statistics*, 1987. doi: 10.1002/9780470316696. [p1]
- D. Ruppert. *Bayesian Data Analysis and MCMC*, pages 531–578. Springer New York, New York, NY, 2011. ISBN 978-1-4419-7787-8. doi: 10.1007/978-1-4419-7787-8_20. URL https://doi.org/10.1007/978-1-4419-7787-8_20. [p11]
- SAS Institute Inc. *Guía del usuario de SAS / STAT, versión 8*. SAS Institute, Cary, NC, 1999. [p2]

- F. Solari, B. Liseo, and D. Sun. Some remarks on bayesian inference for one-way anova models. *Annals of the Institute of Statistical Mathematics*, 60:483–498, 02 2008. doi: 10.1007/s10463-007-0117-5. [p4]
- Stan Development Team. RStan: the R interface to Stan, 2020. URL <http://mc-stan.org/>. R package version 2.21.2. [p10]
- StataCorp. *STATA USER'S GUIDE RELEASE 16*. StataCorp LLC., College Station, TX, 2019. [p2]
- D. van den Bergh, J. van Doorn, M. Marsman, T. Draws, E.-J. van Kesteren, K. Derks, F. Dablander, Q. F. Gronau, Šimon Kucharský, A. R. K. N. Gupta, A. Sarafoglou, J. G. Voelkel, A. Stefan, A. Ly, M. Hinne, D. Matzke, and E.-J. Wagenmakers. A tutorial on conducting and interpreting a bayesian anova in jasp. *L'Année psychologique*, 120, 2020. doi: 10.3917/anpsy1.201.0073. URL <https://www.cairn.info/revue-l-annee-psychologique-2020-1-page-73.htm>. [p2, 5]
- D. Vats and C. Knudson. Revisiting the gelman-rubin diagnostic, 2020. [p11]
- E.-J. Wagenmakers, J. Love, M. Marsman, T. Jamil, A. Ly, J. Verhagen, R. Selker, Q. F. Gronau, D. Dropmann, B. Boutin, F. Meerhoff, P. Knight, A. Raj, E.-J. van Kesteren, J. van Doorn, M. Smíra, S. E. A. Et, D. Matzke, T. de Jong, D. van den Bergh, A. Sarafoglou, H. Steingroever, K. Derks, J. N. Rouder, and R. D. Morey. Bayesian inference for psychology. part ii: Example applications with jasp. *Psychonomic Bulletin & Review*, 25, 2018. doi: 10.3758/s13423-017-1323-7. URL <https://doi.org/10.3758/s13423-017-1323-7>. [p1, 2]
- M. Wedel and C. Dong. Banova: Bayesian analysis of experiments in consumer psychology. *Journal of Consumer Psychology*, 30(1):3–23, 2020. doi: <https://doi.org/10.1002/jcpy.1111>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jcpy.1111>. [p1, 2]
- N. Çelik and B. Şenoğlu. Robust estimation and testing in one-way anova for type ii censored samples: skew normal error terms. *Journal of Statistical Computation and Simulation*, 88(7):1382–1393, 2018. doi: 10.1080/00949655.2018.1433670. URL <https://doi.org/10.1080/00949655.2018.1433670>. [p1]

Johny J. Pambabay-Calero
Faculty of Natural Sciences and Mathematics
ESPOL, Polytechnic University
Guayaquil, Ecuador
<https://orcid.org/0000-0002-9474-9586>
ORCID: 0000-0002-9474-9586
jpambaba@espol.edu.ec

Mauricio J. Rojas-Campuzano
Faculty of Natural Sciences and Mathematics
ESPOL, Polytechnic University
Guayaquil, Ecuador
<https://orcid.org/0000-0001-8000-9432>
ORCID: 0000-0001-8000-9432
maujroja@espol.edu.ec

Sergio A. Bauz-Olvera
Faculty of Natural Sciences and Mathematics
ESPOL, Polytechnic University
Guayaquil, Ecuador
<https://orcid.org/0000-0003-2366-5358>
ORCID: 0000-0003-2366-5358
serabauz@espol.edu.ec

Omar H. Ruiz-Barzola
Faculty of Natural Sciences and Mathematics
ESPOL, Polytechnic University
Guayaquil, Ecuador
<https://orcid.org/0000-0001-8206-1744>
ORCID: 0000-0001-8206-1744
oruiz@espol.edu.ec