

# estimtf: Paquete de R para la estimación de parámetros de múltiples distribuciones de probabilidad y modelos de regresión usando TensorFlow

**Abstract** El lenguaje de programación R provee una amplia variedad de funciones creadas para estimar parámetros de distribuciones y de modelos de regresión usando diversos métodos de estimación. El paquete **estimtf** fue diseñado con el objetivo de encontrar los estimadores de máxima verosimilitud de parámetros distribucionales y de regresión usando TensorFlow, una librería de código abierto para computación numérica. Para alcanzar este objetivo, las principales funciones de **estimtf** implementan un proceso de estimación iterativo en el cual se utilizan los optimizadores incluidos en esta librería para maximizar la función de verosimilitud. Este paquete brinda a los usuarios de R una herramienta que aprovecha las bondades de TensorFlow para la solución de problemas de estimación y además permite hallar los estimadores de parámetros de múltiples distribuciones y modelos de regresión que no están implementados aún en R.

**Palabras clave:** TensorFlow, estimación de parámetros, máxima verosimilitud, algoritmos de optimización

## Introducción

El paquete **estimtf** es un paquete de R creado por [Garcés and Hernández \(2021\)](#) para hallar los estimadores de máxima verosimilitud de parámetros de múltiples distribuciones de probabilidad y modelos de regresión lineal usando TensorFlow.

TensorFlow es una librería de código abierto creada por Google, la cual usa grafos de flujo de datos para llevar a cabo cálculos numéricos ([Galeone, 2019](#)). Esta librería fue creada principalmente para la implementación y desarrollo de modelos de aprendizaje automático, sin embargo, también puede ser utilizada en otros ámbitos como la optimización numérica ([Abadi et al., 2016](#)). Los usuarios del lenguaje de programación R pueden utilizar esta librería a través del paquete **tensorflow** ([Allaire and Tang, 2021](#)).

Uno de los principales componentes de TensorFlow son sus optimizadores que son algoritmos que utilizan optimización numérica basada en gradientes para actualizar los parámetros de una función con el objetivo de minimizarla ([Allaire and Tang, 2021](#)). Para hallar los estimadores de máxima verosimilitud de parámetros distribucionales y de modelos de regresión, las principales funciones del paquete **estimtf** implementan un proceso iterativo en el cual se utiliza un optimizador de TensorFlow para minimizar el negativo de la función de log-verosimilitud. Durante este proceso, los parámetros de interés varían de una iteración a otra a partir de una regla de actualización que depende del optimizador seleccionado. Este proceso continúa hasta que la diferencia entre el valor de los parámetros en la iteración actual y en la iteración inmediatamente anterior es muy pequeña o hasta que se alcance el máximo número de iteraciones.

## Principales funciones del paquete estimtf

A continuación se presentan las principales funciones del paquete **estimtf**. Para utilizar el paquete **estimtf** y estas funciones se recomienda visitar este [repositorio](#) y seguir las instrucciones para su correcta instalación.

```
mle_tf()
```

Esta función permite hallar los estimadores de máxima verosimilitud de los parámetros de múltiples distribuciones de probabilidad usando los optimizadores de TensorFlow. `mle_tf` utiliza el proceso de estimación mencionado anteriormente y requiere que el usuario defina la distribución de interés, los valores iniciales para los parámetros a estimar, el optimizador de TensorFlow que desea utilizar, el máximo número de iteraciones, entre otros.

`mlereg_tf()`

Esta función permite hallar los estimadores de máxima verosimilitud de los parámetros de múltiples modelos de regresión lineal usando los optimizadores de TensorFlow. `mlereg_tf` utiliza el proceso de estimación mencionado anteriormente y requiere que el usuario defina la distribución de la variable respuesta, los valores iniciales para los parámetros a estimar, el optimizador de TensorFlow que desea utilizar, el máximo número de iteraciones, entre otros.

## Ejemplo de uso de la función `mlereg_tf()`

En esta sección se presenta una aplicación en la cual se simulan datos de un modelo de regresión lineal simple tal que  $Y \sim N(\mu, \sigma^2)$  donde  $\mu = \beta_{10} - \beta_{11}x$ ,  $\sigma = \beta_{20}$  y  $x \sim U(-3, 3)$  y se utiliza la función `mlereg_tf()` para estimar  $\beta_{10} = 5$ ,  $\beta_{11} = -2$  y  $\beta_{20} = 3$ .

```
# Cargar el paquete estimtf
library(estimtf)

# Simular datos de un modelo de regresión lineal simple
x <- runif(n = 1000, min = -3, max = 3)
y <- rnorm(n = 1000, mean = 5 - 2 * x, sd = 3)
data <- data.frame(y = y, x = x)

# Usar la función mlereg_tf para estimar los parámetros de interés
mle <- mlereg_tf(ydist = y ~ Normal, formulas = list(mean = ~ x, sd = ~ 1), data = data,
  link_function = NULL, available_distribution = TRUE, initparam = list(mean=0.5, sd=0.5),
  optimizer = "AdamOptimizer", hyperparameters = list(learning_rate=0.1), maxiter = 10000)

# Imprimir resultados
print(mle)

## Loss function convergence, 1611 iterations needed.
## -----
## Distributional parameter: mean
## (Intercept)      x
##      4.989 -2.015
## -----
## Distributional parameter: sd
## (Intercept)
##      2.947
## -----
```

## Referencias

- M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, M. Dean, J. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng. Tensorflow: A system for large-scale machine learning. *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*, 2016. [p1]
- J. Allaire and Y. Tang. *tensorflow: R Interface to "TensorFlow"*, 2021. URL <https://github.com/rstudio/tensorflow>. [p1]
- P. Galeone. *Hands-on neural networks with TensorFlow 2.0: understand TensorFlow, from static graph to eager execution, and design neural networks*. Packt Publishing, 1 edition, 2019. [p1]
- S. Garcés and F. Hernández. *estimtf: Estimation of Distributional and Regression Parameters using TensorFlow*, 2021. URL <https://github.com/SaraGarcesCespedes/estimtf>. R package version 0.1.0. [p1]