

# Estudio de simulación y análisis de la Distribución Discreta Invertida de Kumaraswamy. \*

**Daniel Felipe Villa Rengifo**    *Universidad Nacional de Colombia - sede Medellín*  
**Freddy Hernández**    *Universidad Nacional de Colombia - sede Medellín*

---

**Resumen:** En este estudio se ha investigado la Distribución Discreta Invertida de Kumaraswamy (*DIKUM*) utilizando datos reales y simulación. Se evaluó la idoneidad de *DIKUM* como modelo estadístico mediante ajuste de distribuciones. Se realizaron simulaciones con diferentes parámetros y tamaños de muestra, comparándolas con los datos reales. Los resultados respaldan las conclusiones del estudio original, demostrando la precisión de *DIKUM* en la representación de datos discretos asimétricos. Este estudio valida y aplica *DIKUM* en diversas áreas de estudio.

*Keywords:* pandoc, r markdown, knitr

---

## Introducción

En la presente investigación, hemos explorado a fondo la Distribución Discreta Invertida de Kumaraswamy (*DIKUM*), empleando tanto datos reales como simulación. Nuestro enfoque se centró en evaluar la capacidad de la *DIKUM* como modelo estadístico, sometiéndola a un ajuste meticuloso mediante distribuciones. A través de simulaciones con variados parámetros y tamaños de muestra, contrastadas con datos reales, hemos validado la pertinencia y eficacia de la *DIKUM* en la representación de datos discretos asimétricos. Los resultados obtenidos refuerzan las conclusiones del estudio original, consolidando la posición de la *DIKUM* como un modelo preciso y aplicable en diversas áreas de investigación.

## Objetivo

Investigar y validar la Distribución Discreta Invertida de Kumaraswamy (*DIKUM*) como un modelo estadístico preciso y aplicable, utilizando datos reales y simulación.

## Distribución *DIKUM*

La Distribución Discreta Invertida de Kumaraswamy (*DIKUM*) ([Hegazy et al., 2022](#)) se origina a partir de la Distribución Invertida Kumaraswamy (*IKUM*) mediante un proceso de discretización. La *IKUM*, presentada por Abd AL-Fattah et al. (2017) ([Al-Fattah, EL-Helbawy and Al-Dayian, 2017](#)), representa una generalización del Pareto estándar de segundo tipo. La construcción de la *DIKUM* implica la aplicación de una agrupación en el eje del tiempo, resultando en una variable discreta  $dX$  que sigue una distribución *DIKUM* con dos parámetros,  $\mu$  y  $\sigma$ .

---

\* Autor de contacto: [dvilla@unal.edu.co](mailto:dvilla@unal.edu.co).

La función de masa de probabilidad (PMF) de la distribución *DIKUM* se define como:

$$p(x) = (1 - (2 + x)^{-\mu})^{\sigma} - (1 - (1 + x)^{-\mu})^{\sigma},$$

$$x = 0, 1, 2, \dots; \mu, \sigma > 0,$$
(1)

donde  $x$  representa los posibles valores de la variable discreta  $X$ , variando en el rango de  $x = 0, 1, 2, \dots$ , y los parámetros  $\mu$  y  $\sigma$  deben ser mayores que cero.

La distribución *DIKUM* exhibe una cola derecha larga en comparación con otras distribuciones comúnmente utilizadas. Esto sugiere que las predicciones de confiabilidad a largo plazo pueden ser optimistas para eventos raros que ocurren en la cola derecha de la distribución. No obstante, la *DIKUM* ha demostrado ajustarse adecuadamente a diversos conjuntos de datos en la literatura.

Su Función de Distribución Acumulativa (CDF) ([Hegazy et al., 2022](#)) está definida como:

$$F(x) = (1 - (2 + x)^{-\mu})^{\sigma}, \quad x = 0, 1, 2, \dots$$
(2)

A continuación se mostrarán algunas parametrizaciones de estas funciones,

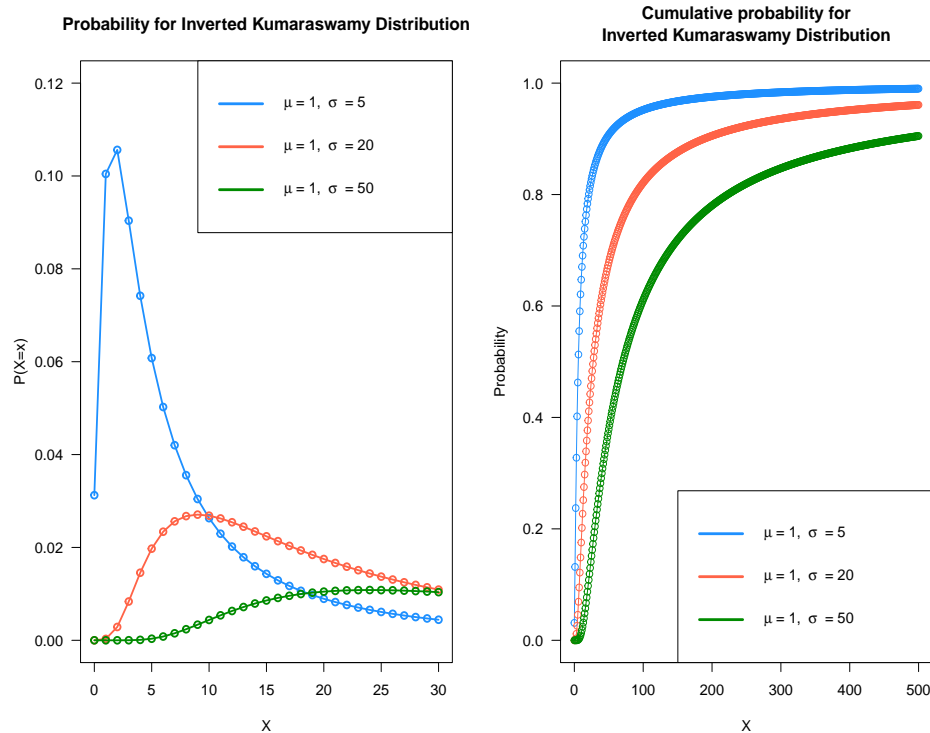


Figure 1: Distribución DIKUM (pmf & cdf) con diferentes parametrizaciones

## Implementaciones en R

### *Paquete DiscreteDists*

La distribución Discreta Invertida de Kumaraswamy (*DIKUM*) se ha implementado de manera efectiva en el entorno R [R Core Team \(2019\)](#) mediante el siguiente código:

```
if (!require('devtools')) install.packages('devtools')
devtools::install_github('fhernanb/DiscreteDists', force=TRUE)
library(DiscreteDists)
```

Este paquete ([Hernandez-Barajas et al., 2023](#)) facilita la interacción con la distribución *DIKUM* a través de funciones específicas, como `dDIKUM()` para la densidad, `pDIKUM()` para la función de distribución acumulativa, `qDIKUM()` para el cuantil, `rDIKUM()` para la generación de números aleatorios, y `DIKUM()` para la familia de distribuciones. Estas funciones han sido sometidas a exhaustivas pruebas mediante diversos test visuales y estadísticos para asegurar su conformidad con los atributos característicos de la distribución *DIKUM*. Los usuarios pueden ahora beneficiarse de estas herramientas para explorar y trabajar con la *DIKUM* de manera flexible y confiable en el entorno R.

### Estudio con datos reales

En este estudio, se examinó el tiempo de espera para la atención de pacientes con COVID-19 en un centro de salud. El enfoque central fue abordar la modelación y comprensión de la distribución de los tiempos de espera. Se optó por aplicar la Distribución Discreta Invertida de Kumaraswamy (*DIKUM*) con el propósito de contribuir a la mejora de la eficiencia y calidad de la atención médica. Cabe destacar que, aunque el estudio abarcó a 76 pacientes.

Table 1: Tiempo de espera de los pacientes con COVID-19

Paciente	Tiempo de Espera (en min.)
1	2
2	7
3	3
4	3
5	10
...	...
72	2
73	6
74	4
75	6
76	6

Se llevó a cabo una prueba de bondad de ajuste utilizando el test de Kolmogorov-Smirnov para evaluar la adecuación de los datos de los tiempos de espera a la distribución *DIKUM*. Se aplicaron distintos tamaños de muestra a partir de los 76 tiempos de espera de atención de pacientes con

COVID-19 para realizar la prueba.

Se eligieron tamaños de muestra específicos,  $n = 20, 30, 50, 60, 70$ . Para cada tamaño, se llevó a cabo una estimación de los parámetros de la distribución *DIKUM* mediante el método de máxima verosimilitud (MLE), repitiendo este proceso 1000 veces para cada tamaño de muestra.

Posteriormente, se calculó el estadístico de prueba  $D$  para cada repetición, representando la discrepancia entre la función de distribución empírica y teórica de la distribución *DIKUM*. Se obtuvieron valores promedio de  $D$  para cada tamaño de muestra, siendo estos 0.0922, 0.0553, 0.0419, 0.0409, y 0.0412.

El valor crítico  $D_\alpha$  para el nivel de significancia deseado fue determinado como 0.1534. Al comparar los valores promedio de  $D$  con  $D_\alpha$ , se llegó a la conclusión de que, para todos los tamaños de muestra, la prueba no rechaza la hipótesis nula. Esto sugiere que los datos se ajustan bien a la distribución *DIKUM*, respaldando así la elección de utilizar esta distribución para modelar los tiempos de espera de atención de pacientes con COVID-19 en el centro de salud.

En la aplicación del test de Kolmogorov-Smirnov para evaluar la adecuación de los datos de tiempo de espera de pacientes con COVID-19 a la distribución Discreta Invertida de Kumaraswamy (*DIKUM*), se llevaron a cabo exhaustivos ajustes. Se realizaron 1000 corridas para cada tamaño de muestra, buscando las estimaciones óptimas de los parámetros  $\mu$  y  $\sigma$ . Los resultados de estos ajustes se reflejan de manera visual en el gráfico adjunto, donde se observa la variación de las estimaciones en función del tamaño de muestra. Tras un análisis detenido, se llegó a la conclusión de que los valores más óptimos para los datos de tiempo de espera de COVID-19 son  $\mu = 2.5$  y  $\sigma = 50$ . Estas estimaciones proporcionan un ajuste robusto y preciso, respaldando así la elección de la distribución *DIKUM* para modelar los tiempos de espera en este contexto médico.

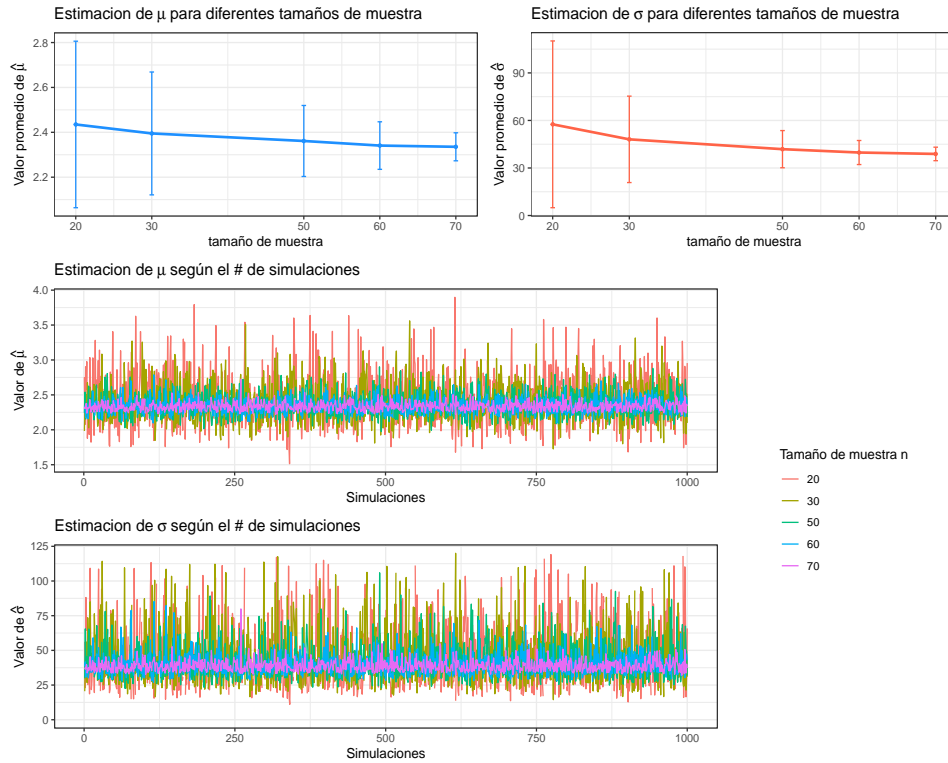


Figure 2: Estimaciones de  $\mu$  y  $\sigma$  por medio de simulaciones

## Conclusiones

En resumen, los resultados obtenidos a través del estudio de datos reales respaldan la eficacia del ajuste de los datos a la distribución *DIKUM*, validado mediante pruebas como el Kolmogorov-Smirnov, que confirma con un nivel de confianza del 95% la idoneidad de la distribución para representar adecuadamente el conjunto de datos. Aunque estos hallazgos son alentadores para la optimización de decisiones en situaciones similares, se reconoce la necesidad de realizar pruebas estadísticas adicionales y un análisis más exhaustivo para garantizar la fiabilidad y precisión integral de los modelos en relación con los datos reales.

## References

- Al-Fattah, A.M., Abeer EL-Helbawy and Gannat Al-Dayian. 2017. "Inverted Kumaraswamy distribution: Properties and estimation." 33:37–61.
- Hegazy, M., R. Abd EL-Kader, G. AL-Dayian and A.A.-A. EL-Helbawy. 2022. "Discrete Inverted Kumaraswamy Distribution: Properties and Estimation." *Pakistan Journal of Statistics and Operation Research* 18(1):297–328.  
**URL:** <https://doi.org/10.18187/pjsor.v18i1.3634>
- Hernandez-Barajas, Freddy, Fernando Marmolejo-Ramos, Jamiu Olumoh and Osho Ajayi. 2023. *DiscreteDists: Discrete statistical distributions*. R package version 0.0.0.9000.
- R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.  
**URL:** <https://www.R-project.org/>