

predfinitepop: Una Librería de R para Hacer Inferencia sobre Totales y Promedios de Poblaciones Finitas usando Métodos Bayesianos No Paramétricos

Objetivo:

La librería “predfinitepop” de R realiza inferencias sobre totales y promedios de poblaciones finitas usando datos muestrales y métodos bayesianos no paramétricos (*species-sampling models*).

Beneficios principales:

1. Versatilidad inferencial al producir inferencias distribucionales de totales y promedios de poblaciones finitas.
2. Genera inferencias sobre totales y promedios de estratos o dominios planeados de una población.
3. Genera inferencias sobre totales y promedios de dominios no planeados, así como de la composición de la población dentro de los dominios no planeados.
4. Genera inferencias robustas, al controlar por la presencia de valores extremos.
5. Versatilidad para generar estadísticas de indicadores económicos basados en muestreo.
6. Contribuye a la construcción de indicadores económicos e institucionales.
7. Aumenta la calidad metodológica de indicadores económicos.
8. Contribuye al desarrollo y divulgación del conocimiento estadístico.

Plataforma de desarrollo:

La librería “predfinitepop” es un módulo desarrollado por empleados del Banco de México, usando a R como lenguaje de desarrollo.¹ Esta librería se desarrolla e implementa dentro de R. Sin embargo, “predfinitepop” se desarrolló independientemente de otras librerías no precargadas en R. La librería “predfinitepop” contiene códigos y rutinas de desarrollo abierto. La librería es auto-contenida y documentada.²

Principales características del producto:

1. Importa a R datos de encuestas con mediciones individuales continuas y variables de segmentación de los datos.
2. Cuenta las frecuencias de las mediciones continuas individuales, respecto a las variables de segmentación de los datos correspondientes.
3. Ajusta modelos de distribuciones con valores extremos para las mediciones individuales de los datos, realiza una comparación predictiva de los modelos ajustados, y selecciona el modelo óptimo predictivo. Las distribuciones consideradas son representativas de la familia de valores extremos.
4. Ajusta un modelo de *species-sampling* para calcular la distribución predictiva de la población, respecto a las variables de segmentación indicadas.
5. Calcula la distribución de los totales de la población, expresados como suma de valores individuales, el cual se calcula como una distribución convolución inducida por los modelos de *species-sampling*.

¹ R es un lenguaje de programación estadística orientada a objetos y de código abierto. URL: <http://www.r-project.org/>

² Véase: Martínez-Ovando et al. (2014) *Package ‘predfinitepop’ to Make Predictive Inference on Totals and Averages of Finite Populations Segmented in Planned and Unplanned Domains*. User’s manual.

6. Ajusta un modelo estadístico de urnas de Pólya para estimar la composición de la población cuando la segmentación de los datos no haya sido planeada para el levantamiento de los datos.
7. Calcula la distribución anidada conjunta de los totales y de la composición de la población cuando la segmentación de los datos no fue planeada en el levantamiento.
8. Calcula la distribución de Monte Carlo de los promedios poblacionales para segmentaciones planeadas y no planeadas de la población.
9. Calcula estimadores puntuales y por intervalos de totales y promedios poblacionales usando datos muestrales, para segmentaciones de la muestra y de la población planeadas en el diseño de la encuesta, y para segmentaciones de la población no considerados en el diseño.