

Análisis de Campos Aleatorios Gaussianos para Gastos en Alimentos para Animales en Puno - ENA 2024

Estudiante Yhack Bryan Aycaya Paco
Docente Dr. Fred Torres Cruz
Curso Estadística Espacial
Institución Universidad Nacional del Altiplano (UNA PUNO)

1. Introducción

Este informe presenta un análisis espacial de los gastos en alimentos para animales (excluyendo alimentos balanceados) en la región de Puno, basado en los datos de la Encuesta Nacional Agropecuaria (ENA) 2024, específicamente la variable P1002B_1C. Se emplean técnicas de Campos Aleatorios Gaussianos (GRF) para modelar la distribución espacial de estos gastos, utilizando coordenadas geográficas y un variograma para capturar la variabilidad espacial.

2. Metodología

Los datos fueron procesados siguiendo los pasos detallados a continuación:

1. **Carga y filtrado de datos:** Se cargaron los datasets de la ENA 2024, filtrando únicamente los registros correspondientes a Puno (CCDD = "21").
2. **Generación de códigos ubigeo:** Se creó una variable ubigeo concatenando CCDD, CCPP y CCDI.
3. **Eliminación de duplicados:** En el dataset de carátula, se eliminaron duplicados, conservando la primera coordenada (LATITUD, LONGITUD) por ubigeo.
4. **Unión de coordenadas:** Se unieron las coordenadas al dataset principal usando ubigeo.
5. **Limpieza de datos:** Se seleccionó la variable P1002B_1C y se eliminaron valores nulos.
6. **Manejo de outliers:** Se identificaron y eliminaron outliers utilizando el rango intercuartílico (IQR).
7. **Eliminación de coordenadas duplicadas:** Se eliminaron registros con coordenadas duplicadas, dejando 88 registros válidos.
8. **Filtrado de coordenadas atípicas:** Se aplicaron límites geográficos para Puno (LONGITUD: [-71.5, -68.5], LATITUD: [-17, -13]).
9. **Transformación a coordenadas proyectadas:** Los datos se transformaron al sistema UTM zona 19S.
10. **Análisis espacial:** Se generó un variograma empírico y se ajustó un modelo esférico para realizar kriging ordinario.

3. Resultados

El proceso de limpieza redujo los datos de 2491 registros iniciales a 88 registros válidos tras eliminar nulos, outliers y coordenadas duplicadas/atípicas. El resumen de la variable P1002B_1C en coordenadas UTM mostró:

- Mínimo: 50
- Mediana: 1500
- Media: 1835
- Máximo: 6631

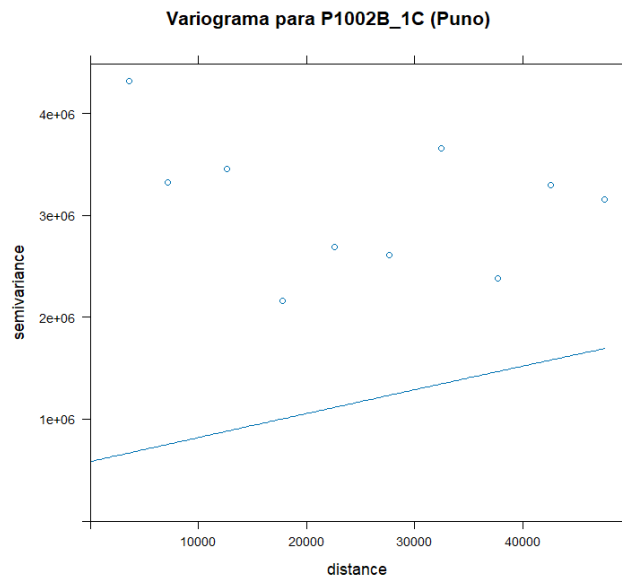


Figura 1: Variograma para P1002B_1C (Puno)

El modelo de variograma ajustado presentó los siguientes parámetros:

- Nugget: 583417.8
- Sill parcial (Sph): 3866439.7
- Rango: 245262.1 m

El kriging ordinario generó predicciones con los siguientes valores:

- Mínimo: 31.81
- Mediana: 1683.09
- Media: 1527.36
- Máximo: 3971.15

4. Conclusiones

El análisis revela una distribución espacial heterogénea de los gastos en alimentos para animales en Puno, con una variabilidad significativa capturada por el modelo de variograma esférico. El kriging ordinario permitió estimar los gastos en áreas no muestreadas, proporcionando una herramienta útil para la planificación agropecuaria en la región. Se recomienda revisar la convergencia del modelo de variograma y considerar datos adicionales para mejorar la robustez del análisis.

Interpolación GRF de P1002B_1C (Puno)

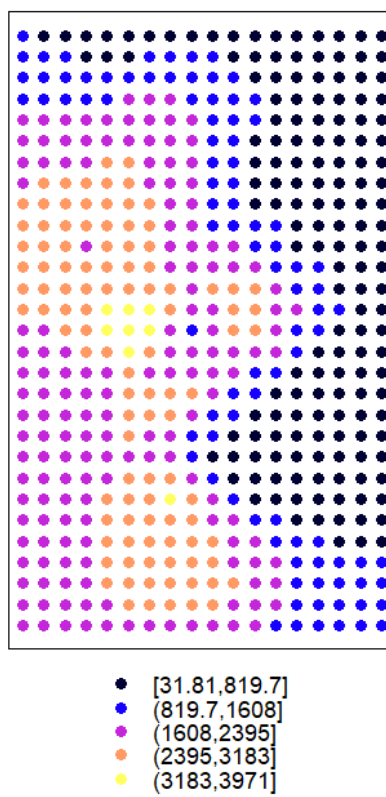


Figura 2: nterpolación GRF de P1002B1C(Puno)