

- 1 Modelos Jerárquicos
- 2 Introducción SAE
- 3 Modelo Fay-Herriot
- 4 Inseguridad alimentaria en Montevideo

- Diseño muestral: procedimiento utilizado para seleccionar unidades de la población
- Estimación de dominio: Estimar una cantidad de interés en un dominio a partir de la información muestral.
- Área pequeña: área o dominio que representa una porción pequeña de la población

Cantidad de interés	Estimación directa
$P_d = \frac{\sum U_d y_{id}}{N_d}$	$\hat{p}_d = \frac{\sum S_d w_{id} y_{id}}{\sum S_d w_{id}}$

U_d dominio poblacional, S_d muestra poblacional, $w_{id} = Pr(y_{id} \in S)^{-1}$ peso muestral.

Área o dominio que representa una porción pequeña de la población

- Cantidad de interés (promedio o proporción) en un área geográfica pequeña (condado, sección censal, barrio, etc)
- Cantidad de interés (promedio o proporción) para un dominio con varias condiciones (hogares de Montevideo con menores de 18, con mujer como jefa de hogar)

Cuando U_d es muy pequeño \hat{p}_d es muy variable o no disponible (si $n_d = 0$)

Fay, R. E. and Herriot R. A. (1979) *Estimates of income for small places: an application of James-Stein procedures to census data*. Journal of the American Statistical Association, 74, 269-277

$$\hat{p}_d = P_d + \epsilon_d, \quad \epsilon_d \sim N(0, V_d)$$

$$P_d = x_d\beta + u_d, \quad u_d \sim N(0, \sigma^2)$$

(p_d, V_d) son conocidas, representan las estimaciones directas de la cantidad de interes y su varianza.

Respuesta = Componente fijo + Componente aleatorio + Error muestral

- Varios niveles de modelos: a nivel del área, a nivel de la unidad, combinando varias fuentes de información
- Varias maneras de estimar los modelos: Frecuencista (BLUP, EBLUP), Bayesiano jerárquico, Bayesiano empírico.

Molina, I. (2019). *Desagregación de datos en encuestas de hogares: metodologías de estimación en áreas pequeñas*. Estudios Estadísticos. CEPAL.

Liu, B., Lahiri, P., & Kalton, G. (2014). *Hierarchical Bayes modeling of survey-weighted small area proportions*. *Survey Methodology*, 40(1), 1-13.

Variable de respuesta: \hat{p}_d la estimación directa de la proporción en el área de interés.

Modelo 2:

$$\hat{p}_d \sim N(P_d, V_d) \quad \text{logit}(P_d) \sim N(x_d\beta, \sigma^2)$$

las varianzas $V_d = \frac{\hat{p}(1-\hat{p})}{n_d} \text{deff}_d$ son conocidas pues provienen del diseño muestral.

Liu, B., Lahiri, P., & Kalton, G. (2014). *Hierarchical Bayes modeling of survey-weighted small area proportions*. *Survey Methodology*, 40(1), 1-13.

Variable de respuesta: \hat{p}_d la *estimación directa* de la proporción en el área de interés.

Modelo 4:

$$\hat{p}_d \sim \text{Beta}(a_d, b_d) \quad \text{logit}(P_d) \sim N(x_d\beta, \sigma^2)$$

donde $a_d = P_d(\frac{n_d}{\text{deff}_d} - 1)$, $b_d = (1 - P_d)(\frac{n_d}{\text{deff}_d} - 1)$ son conocidos.

$$\beta \sim N(0, 3) \quad \sigma \sim N^+(0, 3)$$

- β opera en la escala *logit()*, el 99 % del intervalo $(0, 1)$ se mapea a $(-5, 5)$ aproximadamente.
- $\sigma^2 \sim IG(\epsilon, \epsilon)$ puede ser restrictiva. Otras previas sugeridas pueden ser $Ca^+(0, 1)$ o *Uniforme* $(0, A)$ con A grande.

1 Modelos Jerárquicos

2 Introducción SAE

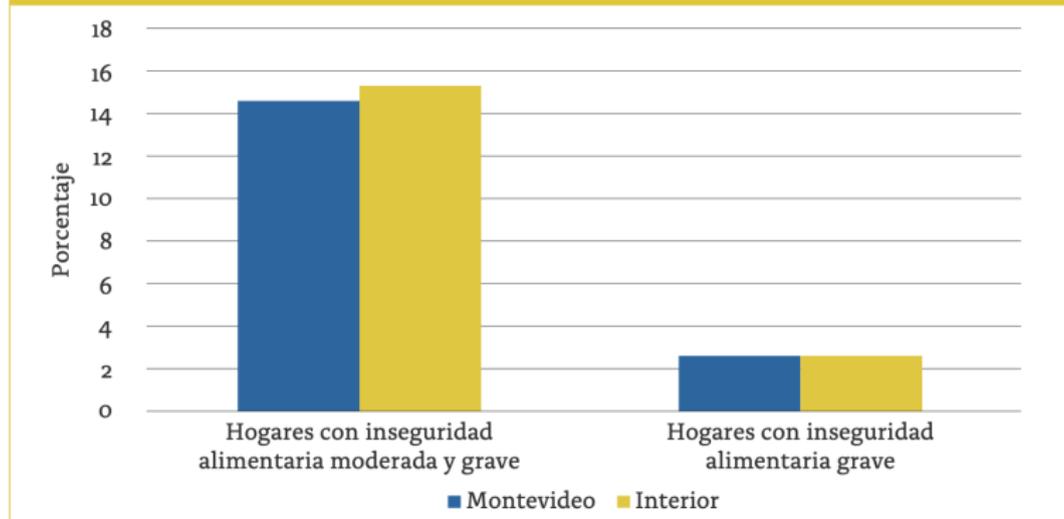
3 Modelo Fay-Herriot

4 Inseguridad alimentaria en Montevideo

¿Qué entendemos por inseguridad alimentaria?

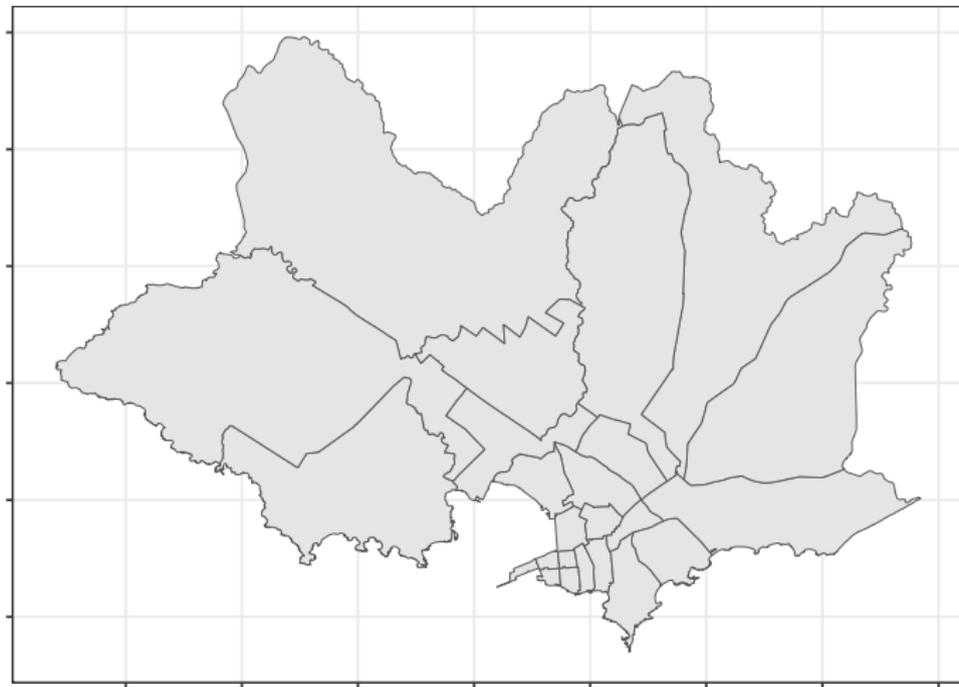


Gráfico 1 - Prevalencia de inseguridad alimentaria moderada y grave en hogares por región (%)



Estimación INE 2022

Secciones de Montevideo



Estimamos el *Modelo 2* en STAN

```
library(rstan)
rstan_options(auto_write = TRUE)

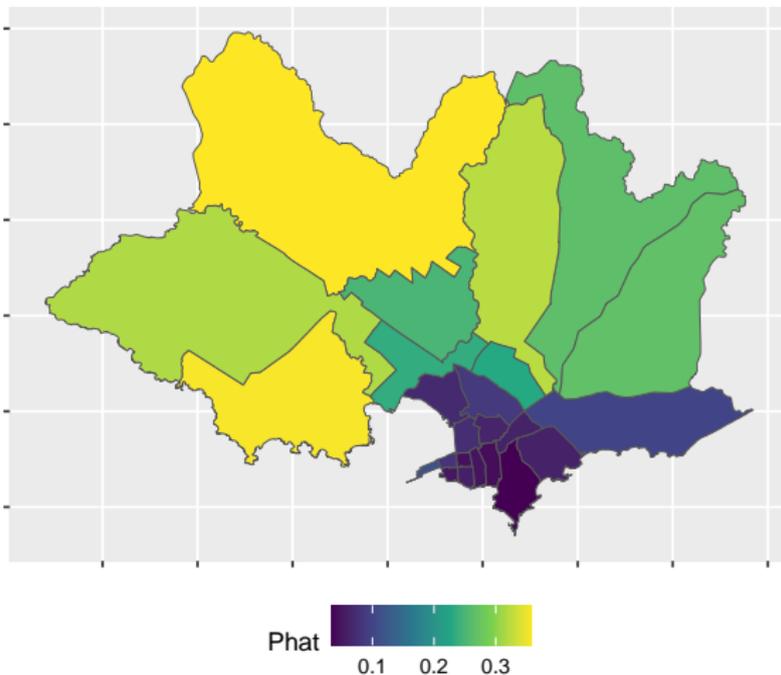
# Armar datos y estimar el modelo
dd <- list(
  S = 24, p=insAlim$iaMoG, SE=insAlim$se, k = 1, X=matrix(1, ncol = 1,
)

fit1 <- stan(file = here('rcode/modelo2.stan'), data=dd, verbose = FALSE)

# Verificar convergencia
# plot(fit1, plotfun='rhat', bins=50)
# plot(fit1, plotfun='ess', bins=50)

# Graficar los P estimados
# plot(fit1, pars = 'P', bins=50, ci_level = 0.5)
```

Inseguridad Alimentaria por secciones



Introducción a la estadística Bayesiana con aplicaciones de estimación en áreas pequeñas usando software STAN

Ignacio Alvarez-Castro Juan José Goyeneche

Instituto de Estadística, Facultad de Ciencias Económicas y Administración, Udelar.

XV Congreso Latinoamericano de Sociedades de Estadística
9 al 13 de Octubre 2023
Santiago de Cali, Colombia