CONSTRUYA 4 CONGLOMERADOS EN LA POBLACIÓN DONDE CADA UNO TIENE 5 ELEMENTOS

Hernan Perci Nuñez Palomino

Ultima edicion 31 Mayo 2021

Table of Contents

# construccion de conglomerados y elementos de la poblacion

# numero de conglomerados  
N <- 4  
# numero de elementos por conglomerado (tamaños iguales)  
M <- 5  
# tabla  
conglomerado <- rep(1:N, M)  
set.seed(123)  
valores <- rnorm(n = N \* M, mean = 0, sd = 1)  
data <- tibble(conglomerado, valores) %>%  
 nest(-conglomerado)  
(unnest(data, -conglomerado))

|  |  |
| --- | --- |
| conglomerado | valores |
| 1 | -0.5604756 |
| 1 | 0.1292877 |
| 1 | -0.6868529 |
| 1 | 0.4007715 |
| 1 | 0.4978505 |
| 2 | -0.2301775 |
| 2 | 1.7150650 |
| 2 | -0.4456620 |
| 2 | 0.1106827 |
| 2 | -1.9666172 |
| 3 | 1.5587083 |
| 3 | 0.4609162 |
| 3 | 1.2240818 |
| 3 | -0.5558411 |
| 3 | 0.7013559 |
| 4 | 0.0705084 |
| 4 | -1.2650612 |
| 4 | 0.3598138 |
| 4 | 1.7869131 |
| 4 | -0.4727914 |

* La media poblacional por elementos es 0.1416238.
* La varianza poblacional es 0.8987739

# muestra de 3 conglomerados en primera etapa

n <- 3  
set.seed(123)  
muestra\_1\_etapa <- data %>%  
 filter(conglomerado %in% sample(1:N, size = n))  
(unnest(muestra\_1\_etapa, -conglomerado))

|  |  |
| --- | --- |
| conglomerado | valores |
| 1 | -0.5604756 |
| 1 | 0.1292877 |
| 1 | -0.6868529 |
| 1 | 0.4007715 |
| 1 | 0.4978505 |
| 3 | 1.5587083 |
| 3 | 0.4609162 |
| 3 | 1.2240818 |
| 3 | -0.5558411 |
| 3 | 0.7013559 |
| 4 | 0.0705084 |
| 4 | -1.2650612 |
| 4 | 0.3598138 |
| 4 | 1.7869131 |
| 4 | -0.4727914 |

# muestra de 3 elementos de cada conglomerado en segunda etapa

m <- 3  
set.seed(123)  
muestra\_2\_etapa <- muestra\_1\_etapa %>%  
 mutate(mi = map(data, ~sample\_n(tbl = ., size = m))) %>%  
 select(conglomerado, mi)  
(unnest(muestra\_2\_etapa, -conglomerado))

|  |  |
| --- | --- |
| conglomerado | valores |
| 1 | -0.6868529 |
| 1 | 0.1292877 |
| 1 | 0.4978505 |
| 3 | 0.4609162 |
| 3 | 0.7013559 |
| 3 | -0.5558411 |
| 4 | 0.3598138 |
| 4 | 0.0705084 |
| 4 | -1.2650612 |

# estimar el total poplacional, media por conglomerados y media por elementos

## estimaciones en segunda etapa

(estimaciones\_2\_etapa <- muestra\_2\_etapa %>%  
 mutate(promedio = map(mi, ~ mean(x = unlist(x = .)))) %>%  
 select(conglomerado, promedio) %>%  
 unnest(promedio) %>%  
 mutate(total = M \* promedio))

|  |  |  |
| --- | --- | --- |
| conglomerado | promedio | total |
| 1 | -0.0199049 | -0.0995244 |
| 3 | 0.2021437 | 1.0107183 |
| 4 | -0.2782463 | -1.3912317 |

## estimaciones en primera etapa

(estimaciones\_1\_etapa <- tibble(media\_conglomerado = c(sum(estimaciones\_2\_etapa$total) / n),   
 total = c(N \* sum(estimaciones\_2\_etapa$total) / n),  
 media\_elemento = c(sum(estimaciones\_2\_etapa$total) / (n \* M))))

|  |  |  |
| --- | --- | --- |
| media\_conglomerado | total | media\_elemento |
| -0.1600126 | -0.6400504 | -0.0320025 |

# varianza de cada estimador

# cuasivarianza poblacional entre conglomerados  
cuasi\_entre <- data %>%  
 mutate(Yi = map(data, ~ sum(.))) %>%  
 unnest(Yi) %>%  
 select(Yi) %>%  
 mutate(S2bi = ((Yi - mean(Yi))^2) / (N - 1)) %>%  
 summarise(S2b = sum(S2bi))  
# cuasivarianza poblacional dentro de conglomerados  
cuasi\_dentro <- data %>%  
 mutate(S2wi = map(data, ~ var(x = .))) %>%  
 unnest(S2wi) %>%  
 select(S2wi)  
# varianzas  
varianza\_total\_estimado <- ((N^2) \* ((1 / n) - (1 / N)) \* cuasi\_entre$S2b) +  
 ((N / n) \* (M^2) \* ((1 / m) - (1 / M)) \* sum(cuasi\_dentro$S2wi))  
varianza\_media\_conglomerado\_estimado <- varianza\_total\_estimado / (N^2)  
varianza\_media\_elemento\_estimado <- varianza\_media\_conglomerado\_estimado / (M^2)

* La varianza del total estimado es 22.2896551.
* La varianza de la media por conglomerado estimado es 1.3931034.
* La varianza de la media por elemento estimado es 0.0557241.

# estimacion de la varianza de cada estimador

# cuasivarianza muestral entre conglomerados  
cuasi\_muestral\_entre <- muestra\_1\_etapa %>%  
 mutate(yi = map(data, ~ sum(.))) %>%  
 unnest(yi) %>%  
 select(yi) %>%  
 mutate(s2bi = ((yi - mean(yi))^2) / (n - 1)) %>%  
 summarise(s2b = sum(s2bi))  
# cuasivarianza muestral dentro de conglomerados  
cuasi\_muestral\_dentro <- muestra\_1\_etapa %>%  
 mutate(s2wi = map(data, ~ var(x = .))) %>%  
 unnest(s2wi) %>%  
 select(s2wi)  
# varianzas estimadas  
estimada\_varianza\_total\_estimado <- ((N^2) \* ((1 / n) - (1 / N)) \* cuasi\_muestral\_entre$s2b) +  
 ((N / n) \* (M^2) \* ((1 / m) - (1 / M)) \* sum(cuasi\_muestral\_dentro$s2wi))  
estimada\_varianza\_media\_conglomerado\_estimado <- estimada\_varianza\_total\_estimado / (N^2)  
estimada\_varianza\_media\_elemento\_estimado <- estimada\_varianza\_media\_conglomerado\_estimado / (M^2)

* La estimada de la varianza del total estimado es 14.8292725.
* La estimada de la varianza de la media por conglomerado estimado es 0.9268295.
* La estimada de la varianza de la media por elemento estimado es 0.0370732.