

prueba_ecocbo_2

```
library(devtools)
library(ggplot2)
library(tibble)
library(dplyr)
library(tidyr)
load_all()
```

Single factor

Pretratamiento de los datos

```
DataPrep <- prep_data(epiDat,
                      type = "counts",
                      Sest.method = "average",
                      cases = 5,
                      N = 100,
                      sites = 3,
                      n = 18,
                      m = 3,
                      k = 50,
                      transformation = "square root",
                      method = "bray",
                      dummy = FALSE,
                      useParallel = TRUE,
                      model = "single.factor")
```

|
|

| 0%

		1%
=		1%
=		2%
==		2%
==		3%
==		4%
===		4%
===		5%
====		5%
====		6%
=====		6%
=====		7%
=====		8%
=====		8%
=====		9%
=====		9%
=====		10%
=====		11%
=====		11%
=====		12%
=====		12%

=====	13%
=====	14%
=====	14%
=====	15%
=====	15%
=====	16%
=====	16%
=====	17%
=====	18%
=====	18%
=====	19%
=====	19%
=====	20%
=====	21%
=====	21%
=====	22%
=====	22%
=====	23%
=====	24%
=====	24%
=====	25%
=====	25%

=====		26%
=====		26%
=====		27%
=====		28%
=====		28%
=====		29%
=====		29%
=====		30%
=====		31%
=====		31%
=====		32%
=====		32%
=====		33%
=====		34%
=====		34%
=====		35%
=====		35%
=====		36%
=====		36%
=====		37%
=====		38%

=====	38%
=====	39%
=====	39%
=====	40%
=====	41%
=====	41%
=====	42%
=====	42%
=====	43%
=====	44%
=====	44%
=====	45%
=====	45%
=====	46%
=====	46%
=====	47%
=====	48%
=====	48%
=====	49%
=====	49%
=====	50%
=====	51%

=====		51%
=====		52%
=====		52%
=====		53%
=====		54%
=====		54%
=====		55%
=====		55%
=====		56%
=====		56%
=====		57%
=====		58%
=====		58%
=====		59%
=====		59%
=====		60%
=====		61%
=====		61%
=====		62%
=====		62%
=====		63%

=====	64%
=====	64%
=====	65%
=====	65%
=====	66%
=====	66%
=====	67%
=====	68%
=====	68%
=====	69%
=====	69%
=====	70%
=====	71%
=====	71%
=====	72%
=====	72%
=====	73%
=====	74%
=====	74%
=====	75%
=====	75%
=====	76%

=====		76%
=====		77%
=====		78%
=====		78%
=====		79%
=====		79%
=====		80%
=====		81%
=====		81%
=====		82%
=====		82%
=====		83%
=====		84%
=====		84%
=====		85%
=====		85%
=====		86%
=====		86%
=====		87%
=====		88%
=====		88%

=====	89%
=====	89%
=====	90%
=====	91%
=====	91%
=====	92%
=====	92%
=====	93%
=====	94%
=====	94%
=====	95%
=====	95%
=====	96%
=====	96%
=====	97%
=====	98%
=====	98%
=====	99%
=====	99%
=====	100%

```
head(DataPrep$Results)
```

	dat.sim	k	m	n	pseudoFH0	pseudoFHa	MSR
[1,]	1	1	3	2	0.5842887	1.631776	0.2530066
[2,]	1	2	3	2	1.0768733	1.069626	0.2805245
[3,]	1	3	3	2	1.0292679	1.987974	0.2226186
[4,]	1	4	3	2	0.9309400	2.413558	0.2157747
[5,]	1	5	3	2	0.8223219	1.593329	0.2852602
[6,]	1	6	3	2	1.0353506	2.154793	0.2255597

Componentes de variación

```
CompVar <- scompvar(DataPrep)
CompVar
```

	Source	Est.var.comp
1	Residual	0.2716559

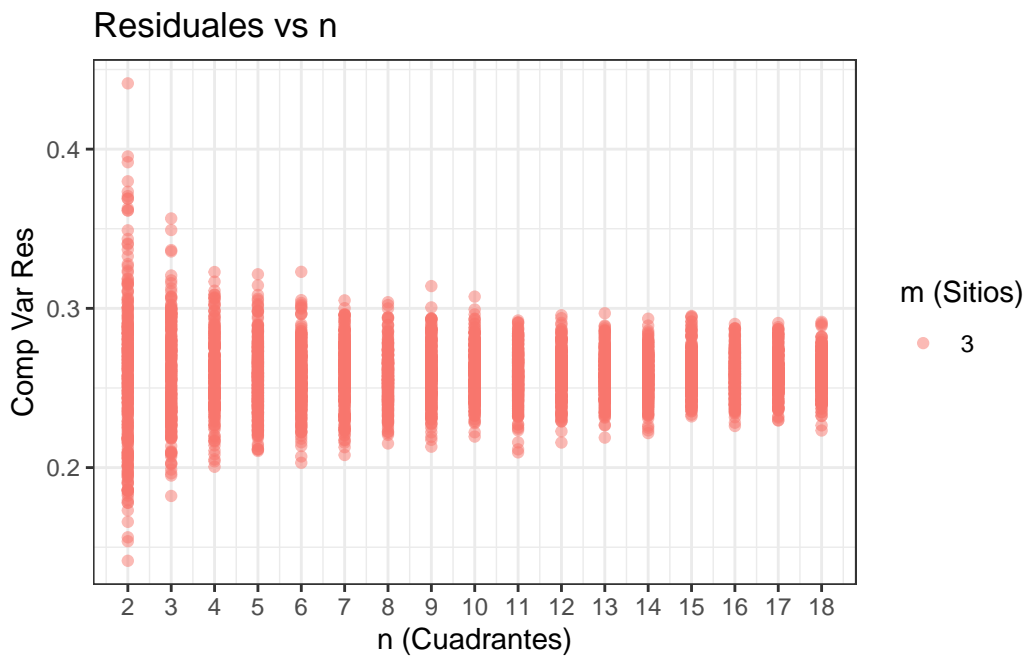
```
# Este valor debe ser igual al máximo de los promedios de cada simulación
CompVarCalculado <- DataPrep$Results |>
  as.data.frame() |>
  group_by(dat.sim) |>
  summarise(MSR = mean(MSR)) |>
  summarise(CompVarCalculado = max(MSR))
CompVarCalculado
```

```
# A tibble: 1 x 1
  CompVarCalculado
          <dbl>
1             0.272
```

Comportamiento de cuadrados medios

Se puede observar como la estimación de componente de variación de los residuales se va haciendo más específica conforme aumenta el número de muestras.

```
ggplot(DataPrep$Results, aes(x = n, y = MSR, color = factor(m))) +
  geom_point(alpha = 0.5) +
  scale_x_continuous(breaks = seq(2, 18, 1)) +
  labs(
    title = "Residuales vs n",
    x = "n (Cuadrantes)",
    y = "Comp Var Res",
    color = "m (Sitios)"
  ) +
  theme_bw()
```



Nested symmetric

Pretratamiento de los datos

```
DataPrep <- prep_data(SSP::epibionts,
  type = "counts",
  Sest.method = "average",
  cases = 5,
  N = 100,
  sites = 3,
```

```

n = 18,
m = 3,
k = 50,
transformation = "square root",
method = "bray",
dummy = FALSE,
useParallel = TRUE,
model = "nested.symmetric")

```

		0%
		1%
=		1%
=		2%
==		2%
==		3%
==		4%
===		4%
===		5%
====		5%
====		6%
=====		6%
=====		7%
=====		8%
=====		8%
=====		9%

=====		9%
=====		10%
=====		11%
=====		11%
=====		12%
=====		12%
=====		13%
=====		14%
=====		14%
=====		15%
=====		15%
=====		16%
=====		16%
=====		17%
=====		18%
=====		18%
=====		19%
=====		19%
=====		20%
=====		21%
=====		21%

=====	22%
=====	22%
=====	23%
=====	24%
=====	24%
=====	25%
=====	25%
=====	26%
=====	26%
=====	27%
=====	28%
=====	28%
=====	29%
=====	29%
=====	30%
=====	31%
=====	31%
=====	32%
=====	32%
=====	33%
=====	34%
=====	34%

=====		35%
=====		35%
=====		36%
=====		36%
=====		37%
=====		38%
=====		38%
=====		39%
=====		39%
=====		40%
=====		41%
=====		41%
=====		42%
=====		42%
=====		43%
=====		44%
=====		44%
=====		45%
=====		45%
=====		46%
=====		46%

=====	47%
=====	48%
=====	48%
=====	49%
=====	49%
=====	50%
=====	51%
=====	51%
=====	52%
=====	52%
=====	53%
=====	54%
=====	54%
=====	55%
=====	55%
=====	56%
=====	56%
=====	57%
=====	58%
=====	58%
=====	59%
=====	59%

	=====	60%
	=====	61%
	=====	61%
	=====	62%
	=====	62%
	=====	63%
	=====	64%
	=====	64%
	=====	65%
	=====	65%
	=====	66%
	=====	66%
	=====	67%
	=====	68%
	=====	68%
	=====	69%
	=====	69%
	=====	70%
	=====	71%
	=====	71%
	=====	72%

=====	72%
=====	73%
=====	74%
=====	74%
=====	75%
=====	75%
=====	76%
=====	76%
=====	77%
=====	78%
=====	78%
=====	79%
=====	79%
=====	80%
=====	81%
=====	81%
=====	82%
=====	82%
=====	83%
=====	84%
=====	84%
=====	85%

	=====	85%
	=====	86%
	=====	86%
	=====	87%
	=====	88%
	=====	88%
	=====	89%
	=====	89%
	=====	90%
	=====	91%
	=====	91%
	=====	92%
	=====	92%
	=====	93%
	=====	94%
	=====	94%
	=====	95%
	=====	95%
	=====	96%
	=====	96%
	=====	97%

```

===== | 98%
|
===== | 98%
|
===== | 99%
|
===== | 99%
|
===== | 100%

```

```
head(DataPrep$Results)
```

	dat.sim	k	m	n	pseudoFH0	pseudoFHa	MSB(A)	MSR
[1,]	1	1	2	2	1.883005	4.595459	0.13604237	0.1678797
[2,]	1	2	2	2	2.094208	3.686394	0.12004485	0.1766198
[3,]	1	3	2	2	1.437205	4.930616	0.07551901	0.1691250
[4,]	1	4	2	2	2.038578	3.614853	0.13360841	0.1456161
[5,]	1	5	2	2	2.079049	5.535402	0.11318552	0.1507316
[6,]	1	6	2	2	1.976757	2.541245	0.15143786	0.1633246

Componentes de variación

```
CompVar <- scompvar(DataPrep)
CompVar
```

	Source	Est.var.comp
1	B(A)	0.0025946
2	Residual	0.2490417

```
# Este valor debe ser igual al máximo de los promedios de cada simulación
```

```
scompvarCalc <- tribble(~name, ~CompVar,
                        "B(A)", NA,
                        "Res", NA)
```

```
CompVarResCalculado <- DataPrep$Results |>
  as.data.frame() |>
  filter(m == 3 & n == 18) |>
  group_by(dat.sim) |>
  summarise(MSR = mean(MSR)) |>
```

```

    summarise(CompVarCalculado = max(MSR))

CompVarBACalculado <- DataPrep$Results |>
  as.data.frame() |>
  filter(m == 3 & n == 18) |>
  group_by(dat.sim) |>
  summarise(MSBA = mean(`MSB(A)`) |>
  summarise(CompVarCalculado = max(MSBA))
CompVarBACalculado <- (CompVarBACalculado - CompVarResCalculado) / 18

scompvarCalc[1,2] <- CompVarBACalculado
scompvarCalc[2,2] <- CompVarResCalculado

scompvarCalc

```

```

# A tibble: 2 x 2
  name   CompVar
  <chr>   <dbl>
1 B(A)  0.00259
2 Res   0.249

```

Comportamiento de cuadrados medios

En los cálculos correspondientes a `scompvar` y demás, la función tomará por default los valores máximos de `n` y `m` para aprovechar el mayor número de grados de libertad posibles. En estas gráficas se incluyen los comportamientos para todos los valores calculados de `n` y `m` para poder observar la operación de las función `prep_data`.

En las primeras dos gráficas se observa el *funnel* para el cálculo de componente de variación residual con respecto al número de muestras `n`, y un comportamiento homogéneo para el cálculo de cuadrados medios anidados $B(A)$ con respecto al número de muestras `n`.

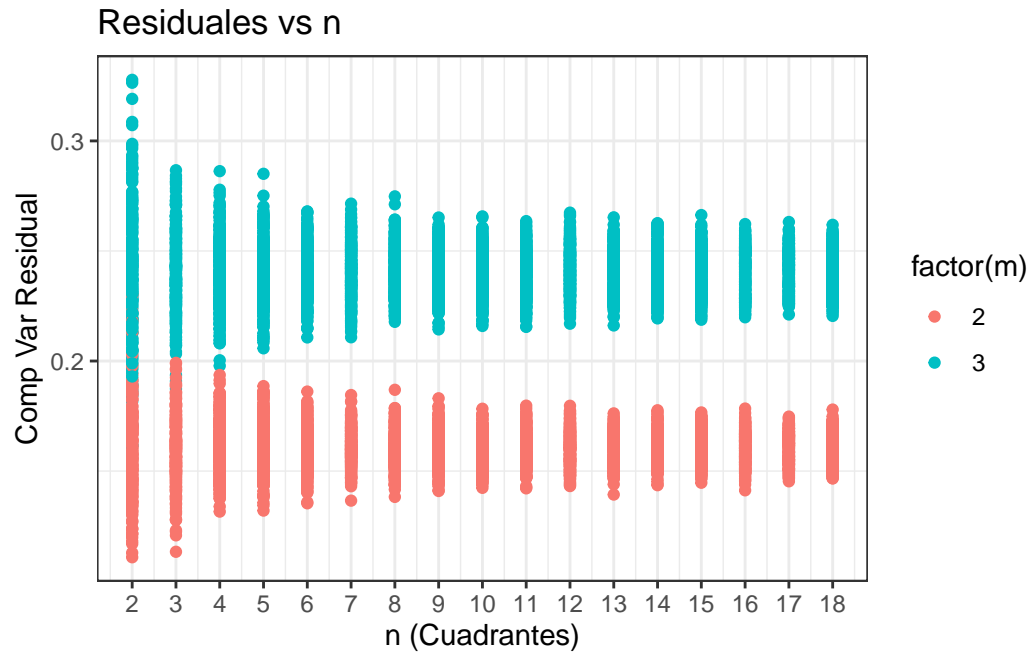
```

# se añade columna de componentes de variación para componente anidado
Resultados <- as.data.frame(DataPrep$Results) |>
  mutate(CompVarBA = (`MSB(A)` - MSR) / n)

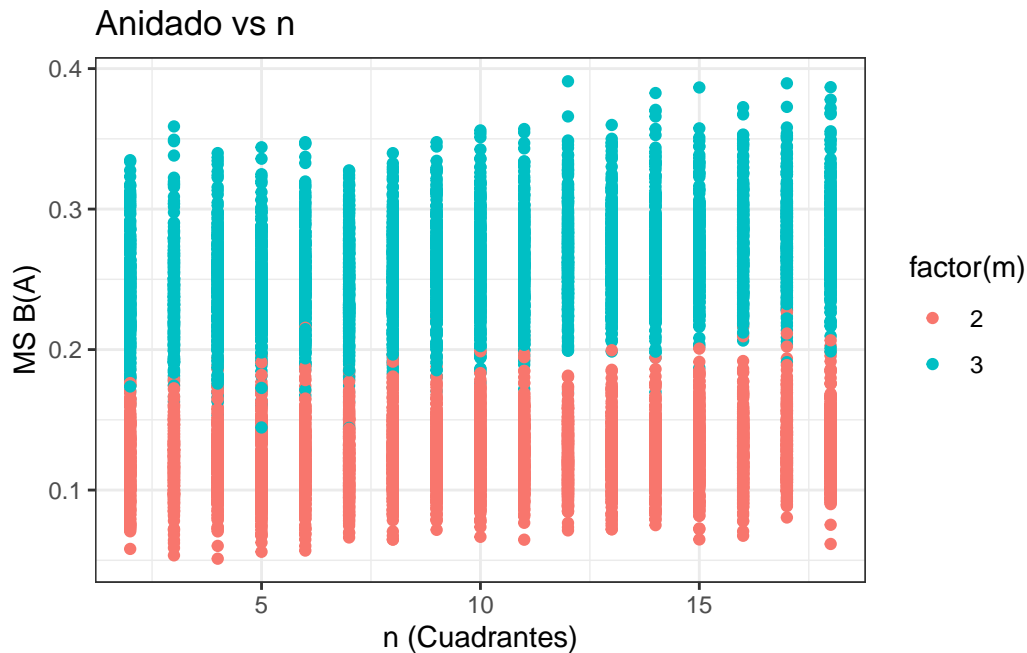
ggplot(Resultados, aes(x = n, y = MSR, color = factor(m))) +
  geom_point() +
  # geom_jitter(height=0)+
  scale_x_continuous(breaks = seq(2, 18, 1)) +
  scale_y_continuous(breaks = seq(0.0, 0.5, 0.1)) +

```

```
labs(
  title = "Residuales vs n",
  x = "n (Cuadrantes)",
  y = "Comp Var Residual",
) +
theme_bw()
```



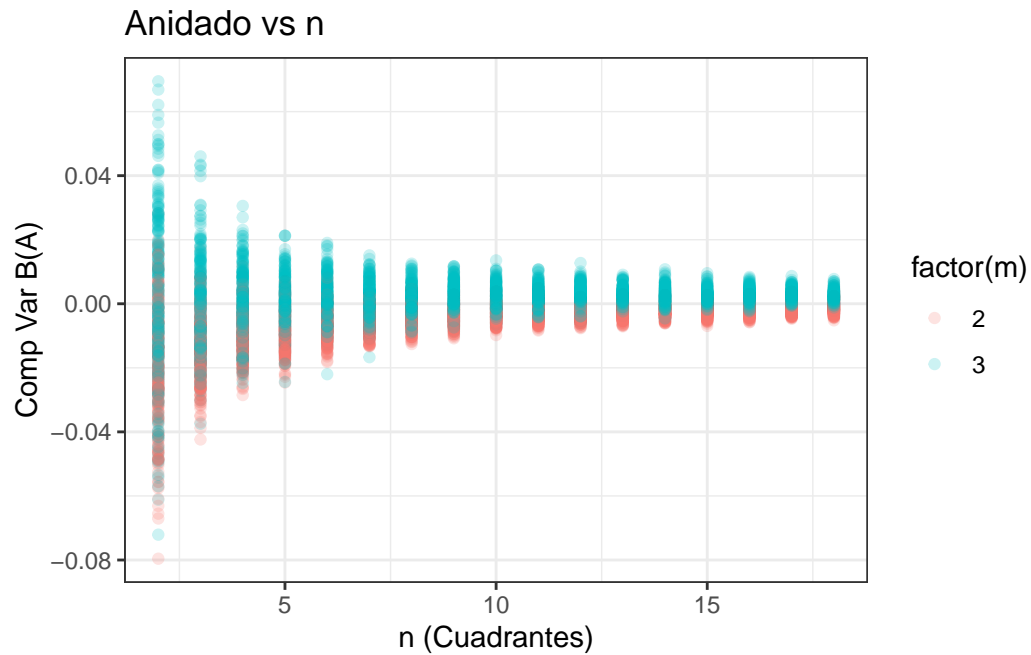
```
ggplot(Resultados, aes(x = n, y = `MSB(A)`, color = factor(m)))+
  geom_point()+
  scale_y_continuous(breaks = seq(0.0, 0.5, 0.1)) +
  labs(
    title = "Anidado vs n",
    x = "n (Cuadrantes)",
    y = "MS B(A)",
  ) +
  theme_bw()
```



A continuación se observan los gráficos para el componente de variación anidado $B(A)$ que tiene forma de *funnel* al igual que el componente de variación residual.

Esto se debe a que el cálculo del componente de variación anidado incluye una división con un denominador que se va haciendo progresivamente mayor, por lo que el resultado tendrá que ser más pequeño cada vez. La última gráfica presenta esta matemática de manera visual, se resta $MS_{B(A)} - MS_R$ y el resultado se divide por un n cada vez mayor, por lo que el componente de variación se va concentrando en valores cada vez más cercanos a cero.

```
ggplot(Resultados, aes(x = n, y = CompVarBA, color = factor(m)))+
  geom_point(alpha = 0.2)+
  labs(
    title = "Anidado vs n",
    x = "n (Cuadrantes)",
    y = "Comp Var B(A)"
  ) +
  theme_bw()
```



```
Resultados |>
  pivot_longer(cols=c("CompVarBA", `MSB(A)`, "MSR"),
               names_to="CompVar",
               values_to="Value") |>
  ggplot(aes(x = n, y = Value, color = CompVar, shape = factor(m)))+
  geom_point(alpha = 0.5)+
  scale_color_manual(values = c("orange", "chocolate", "aquamarine"))+
  labs(title = "Anidado vs n",
       x = "n (Cuadrantes)",
       y = "Componentes",
       color = "Componentes")+
  theme_bw()
```