Análisis del manejo de gatos asilvestrados en isla Martín García

Cliente: Ian Barbe

2023-02-10

# Datos con factores aleatorios

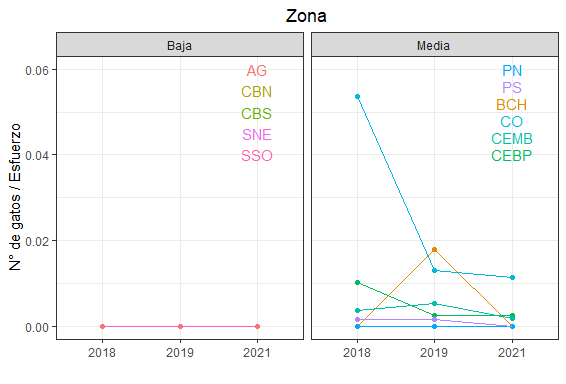
Cuando trabajamos con datos en los que hay una variable que agrupa a las réplicas (en este caso, las transectas agrupan réplicas temporales porque se trata de un diseño de medidas repetidas), lo primero que hacemos es explorar el patrón que se cumple en cada una de las réplicas. Esto lo hacemos porque, cuando corramos un modelo mixto, lo que va a hacer el modelo es **promediar las tendencias** observadas en las transectas de la misma zona. Si las tendencias en las transectas son parecidas, la variabilidad aportada por el factor aleatorio va a ser baja, pero si cada transecta muestra una cosa distinta puede llegar a ocurrir que la variabilidad sea tanta que el efecto en el tiempo se vuelva despreciable.

## Análisis de abundancia de gatos

En la zona Baja no se registraron gatos en ninguna transecta ni en ninguno de los años de muestreo. Acá lo que tenemos es un caso claro en donde los datos **no tienen desviación estándar**. Lo problemático no es que haya muchos ceros, sino que todas las transectas tienen el mismo valor. Tratar de modelar esto va a dar como resultado intervalos de confianza gigantes y no muy confiables en las estimaciones.

Por otro lado, en la zona Media se observaron tendencias temporales diferentes en cada una de las transectas registradas. Mientras que en algunas se mantuvo constante la abundancia de gatos en el tiempo, en otras se observaron patrones de aumento y disminución en diferentes intervalos de tiempo. Este caso sí se puede modelar con un modelo mixto, porque acá sí observamos cambios en el tiempo.

En el caso de que uno quisiera comparar las abundancias entre la zona baja y la zona media, lo que se podría hacer es un no paramétrico que permita la incorporación de efectos aleatorios, como el test de Friedman. La desventaja de este test es que sólo permite evaluar *un efecto fijo* y *un efecto aleatorio*. Sin embargo, con este test se podría responder a la siguiente pregunta: ¿La abundancia de gatos fue **consistentemente mayor** en la zona media en relación a la zona baja **en todos los años de muestreo**?

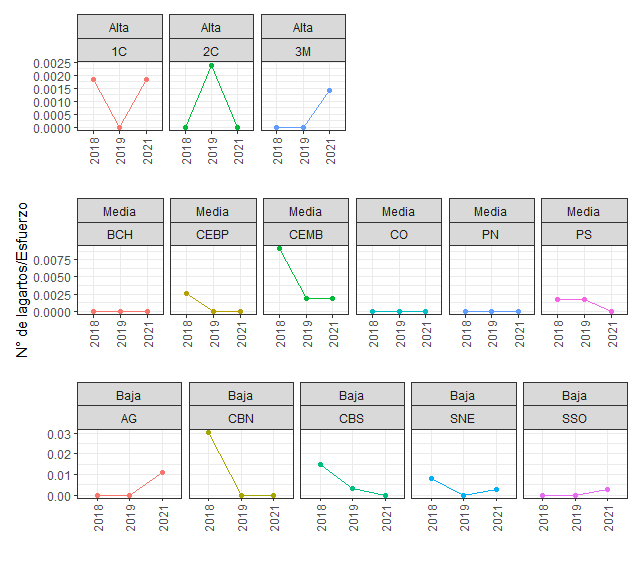


**Figura** **1:** Abundancia de gatos por unidad de esfuerzo registrada para cada transecta (agrupada por zona) en los distintos momentos de muestreo.

## Análisis de abundancia de lagartos

En la zona Alta, se observó una gran variabilidad *intertransecta*. De hecho, en 1C y 2C se notaron patrones temporales opuestos en cuanto a abundancia de lagartos. Esto puede dar a lugar a que el efecto del tiempo se vuelva despreciable. Por su parte, en la zona Media tenemos tres transectas (BCH, CO y PN) en las que no se registraron lagartos (y por lo tanto, no hay variación en el tiempo), mientras que en las otras tres (CEBP, CEMB y PS) se vio un comportamiento de disminución de la abundancia a través del tiempo. Por último, en la zona Baja se observaron tendencias temporales distintas para cada transecta.

Este caso sí se puede modelar completamente con modelos mixtos. Hay que tener en cuenta igual **qué es lo que se quiere buscar** con estos modelos. Por ejemplo, si yo quisiera buscar *el mismo patrón temporal* en todas las zonas, juntaría toooodos los datos en un único modelo, incorporando la zona como factor aleatorio. Ahora, si me estoy dando cuenta de que en cada zona pareciera estar pasando algo distinto, tal vez me conviene subsetear mis datos y armar un modelo para cada zona.



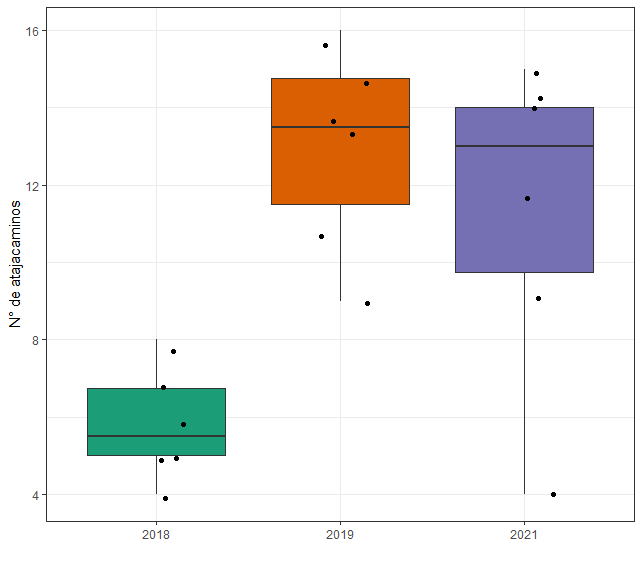
**Figura** **2:** Abundancia de lagartos por unidad de esfuerzo registrada para cada zona y transecta en los distintos momentos de muestreo.

# Datos sin factor aleatorio

Por una cuestión de “cosas que pasan” después de muestrear, hay dos set de datos que no responden a un diseño de medidas repetidas porque ya no hay un factor aleatorio (transecta) que me esté agrupando los datos. Lo bueno de esto, es que simplifica un poco el análisis, ya que la pregunta se reduce a ¿Cómo cambia la abundancia en el tiempo? Esto se puede responder con un modelo de distribución Poisson o Binomial Negativa, o de hecho hasta con un Kruskal-Wallis.

## Análisis de abundancia de atajacaminos

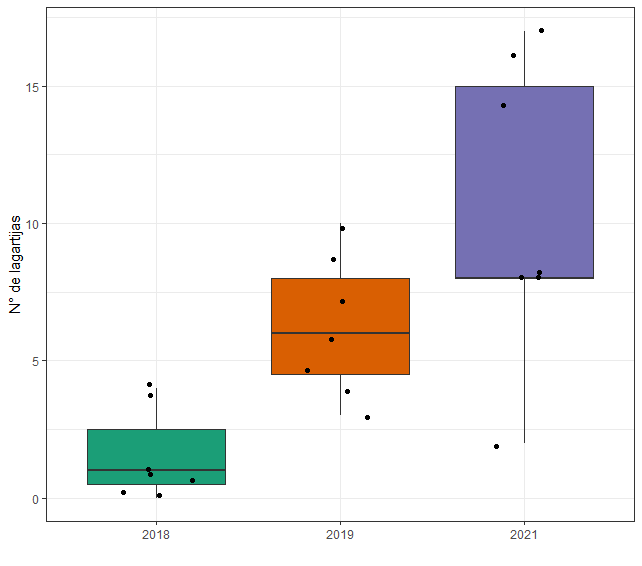
En el gráfico de cajas se pudo observar un claro aumento en la abundancia de atajacaminos en el intervalo 2018 - 2019, mientras que no parece haber cambios en la abundancia en el intervalo 2019 - 2021. Si modelamos esto con un GLM vamos a poder hablar del tamaño, la dirección y la significancia del cambio en el tiempo de la abundancia.



**Figura** **3:** Abundancia de atajacaminos registrada en la transecta ‘Pista’ en cada momento de muestreo.

## Análisis de abundancia de lagartijas

En el gráfico de cajas se puede observar un claro aumento en la abundancia de lagartijas en los tres años de muestreo. Este caso se analiza de la misma manera que el anterior.



**Figura** **4:** Abundancia de lagartijas registrada en la transecta ‘AG’ en cada momento de muestreo.

# Conclusiones del exploratorio

* Para el caso de la abundancia de gatos, es preferible analizar lo que ocurre en la zona Baja y en la zona Media por separado. La zona Baja no tiene observaciones distintas de cero, por lo cual no tiene sentido modelar esos datos. Si yo intentara ingresar esos datos a un modelo es probable que obtenga valores muy raros de estimadores y de errores estándar, porque justamente, los datos no tienen desviación estandar. Por su parte, la zona Media podría modelarse con una distribución poisson/binomial negativa en donde se debería incluir: un offset para tener en cuenta el esfuerzo, un factor aleatorio para incluir el efecto de las transectas y un efecto fijo para analizar diferencias entre años.
* Si se quisiera comparar la abundancia de gatos en la zona Media y la zona Baja, se podría recurrir a un test de Friedman.
* En el caso de los lagartos, se podría construir un modelo mixto en donde las transectas estén anidadas dentro de cada zona. Esto permitiría evaluar si existe un patrón temporal en la abundancia de lagartos que se **mantenga en todas las zonas analizadas**. En caso contrario, si se quisiera buscar patrones distintos en cada zona, podría pensarse realizar un modelo individual para cada una de ellas.
* Con respecto a los atajacaminos y las lagartijas, puede correrse un GLM con distribución binomial negativa / Poisson con el tiempo como único efecto fijo.