

```
#####  
#### Biometria2 - 1er Parcial - 28.9.223 ####  
#####
```

Nombre y apellido:

Turno de Trabajos Practicos:
[Soledad = A / José = B / Nicolás = C]

Correo electronico:

```
#####  
### Antes de comenzar:  
#####
```

Lea atentamente el enunciado y responda lo que se pregunta en el script con comentarios '#'.

Recuerde que su script debe correr. Esto significa que solo con el script y la base de datos en el mismo directorio el docente que corrija su examen debería poder ejecutar y visualizar todos los resultados sin necesidad de agregar o eliminar líneas de código.

Comente el código con #. Esto significa que debe explicitar para que incluye cada línea o grupo de líneas (según corresponda) de comando (ej. si lo hizo para evaluar un supuesto, para ajustar un modelo, para recodificar una variable, para generar un gráfico, etc.).

Para aprobar el examen, además de la correcta resolución del ejercicio, tenga en cuenta que:

- # El script debe correr correctamente
- # No deben faltar comentarios en relación a las líneas de código que decida utilizar
- # Debe existir coherencia interna entre comandos y comentarios.
- # Las conclusiones deben estar fundamentadas

```
### *****  
# El examen es INDIVIDUAL #  
### *****
```

```
### *****  
# Vaya guardando su Script !!!!  
# Apellido_Turno  
# Ej: "Perez_Turno_A"  
# Una vez finalizado, subir al campus  
### *****
```

```
#####
```

Enunciado

#####

#####

Enunciado

#####

La teoría de forrajeo se centra en las decisiones comportamentales que toman los animales para optimizar el esfuerzo dedicado a la búsqueda y explotación de recursos, la exposición al riesgo de depredación y las preferencias dietarias. Estas decisiones conductuales pueden traducirse en patrones de distribución y abundancia. Un proceso clave que vincula los movimientos individuales con la distribución de las especies es la formación y el mantenimiento de las áreas de acción (HR, por su siglas en inglés, home range), es decir, el espacio utilizado para satisfacer las necesidades vitales.

El tamaño del HR refleja una compensación entre los costos y beneficios asociados a la adquisición de recursos, como el alimento, durante un periodo de tiempo determinado. Un determinante clave del tamaño del HR es la disponibilidad de recursos forrajeros, ya que el tamaño del HR disminuye en los entornos más productivos.

Con el fin de estudiar factores que determinan el tamaño del HR, se realizó un estudio en *Canis lupus* (lobo) en Canadá. Esta especie se alimenta principalmente de animales silvestres pero, ante la ausencia de estos, puede alimentarse de animales de cría. Se estudió el HR en 4 ambientes con diferente uso (A1: "bosque nativo", A2: "forestación", A3: "ganadería" y A4: "pastizal") en zonas con nieve y zonas sin nieve. Los investigadores predicen que las zonas con nieve tendrán menor productividad de recursos alimenticios y por lo tanto menor cantidad de presas disponibles, por lo que los individuos de las manadas de esas zonas deberán tener un mayor HR. Sin embargo, sospechan que la disponibilidad de animales de cría puede atenuar estas diferencias.

En cada tipo de ambiente y cada condición de nieve se seleccionaron en forma aleatoria 3 manadas de *C. lupus* y se determinó el HR diario (en km²) de 4 de sus individuos, elegidos al azar entre todos los integrantes de la manada.

Datos en el archivo datos.csv., contiene 96 filas y 5 columnas.

En base al enunciado (INTERCALE LAS RESPUESTAS LUEGO DE CADA PREGUNTA),

Pregunta 1

#####

1 a) Complete el diccionario de variables indicando sus clases. Para las variables explicativas declare además su condición de fijas o aleatorias, cruzadas (con) o anidadas (en) y la cantidad de réplicas.

#Diccionario de variables

Tipo de ambiente: Tipo de ambiente ("bosque nativo", "forestación", "ganadería" y "pastizal"), character/factor. Fija, cualitativa 4 niveles (cruzada con condicion de nieve).

Condición de nieve: estado de la zona (sin nieve y con nieve), character/factor. Fija, cualitativa 2 niveles (cruzadas con tipo de ambiente)

Manada: Identificador único de cada manada, integer, pero deben ponerla como factor. Aleatoria, anidada en la interaccion tipo de ambiente y condicion de nieve.

Individuo: Identificador único de cada individuo, integer.

HR: Home Range (km2), numeric, variable respuesta.

1 b) Escriba el modelo en parámetros (junto con subíndices y supuestos distribucionales), indicando el significado de cada término en contexto (para las letras griegas utilice palabras). Si existiera alguna/s variable/s de efectos aleatorios incorpórela/s de forma implícita.

$HR_{ijkl} = \mu + \alpha_i + \beta_j + \alpha_i * \beta_j + M_{k(ij)} + \epsilon_{ijkl}$

i= 1:4, nivel de tipo de ambiente

j= 1:2, nivel de condicion de nieve

k= 1:3, nivel de manada

l= 1:4, individuos

$M_k(ij) \sim NID(0, \sigma^2_{manadas})$

$\epsilon_{ijkl} \sim NID(0, \sigma^2)$

HR ijkl es el tamaño del home range de cada lobo

μ es la media general

α_i es el efecto medio del i-esimo nivel del tipo de ambiente para una manada típica

β_j es el efecto del j-esimo nivel de la condicion de nieve# $\alpha_i * \beta_j$ es el efecto de la interaccion ij

$M_k(ij)$ es el efecto aleatorio de la manada k-esima anidada en la combinación ambiente-nieve ij

ϵ_{ijkl} es el error aleatorio

Pregunta 2

#####

Realice un gráfico de perfiles en donde se muestre la relación entre el HR y la condición de nieve, discriminando por tipo de ambiente. Escriba una frase describiendo dicha relación.

library(car)

library(ggplot2)

library(lme4)

library(nlme)

```

library(lmerTest)
library(dplyr)

datos <- read.csv(file = "/home/jose/Documents/materias/biome2/2023/parcial/datos.csv")
head(datos)
str(datos)
datos$individuos <- as.factor(datos$individuos)
datos$manada <- as.factor(datos$manada)

datos %>%
  group_by(nieve, areas, manada) %>%
  summarise(n = n())

ggplot(datos, aes(nieve, HR, color = areas)) +
  geom_jitter(width = 0.05, height = 0, size = 6) +
  labs(y = "Home range (km2)", x = "Condición de Nieve")

```

Se observa menor tamaño del home range de los lobos en zonas sin nieve, a excepción del ambiente 3 en el cual los lobos parecen mantener el mismo tamaño del home range tanto en ausencia o presencia de nieve, en el cual sí hay disponibilidad de animales de cría. Esto sugiere una posible interacción entre los efectos de condición de nieve y tipo de ambiente.

Pregunta 3
#####

Ajuste el modelo propuesto por los investigadores. Evalúe los supuestos de dicho modelo y concluya, si corresponde, en base a:

```

m1 <- lmer(HR ~ nieve * areas + (1 | manada), data = datos)
summary(m1)
anova(m1)

```

```

# 3.1- normalidad
# con performance
library(performance)
check_normality(m1)

```

```

#De manera tradicional
e <- residuals(m1, type = "pearson")
# QQ plot
qqPlot(e, main = "QQ Plot residuos")

```

```

# Prueba de normalidad (Shapiro-Wilk test of normality)
shapiro.test(e)

```

No hay evidencias de que los errores no tengan distribución normal (p=...)

```
# b) de los alfai
alfai <- ranef(m1)$manada$(Intercept)"
qqPlot(alfai, main = "QQ Plot ef aleat")
# Prueba de normalidad (Shapiro-Wilk test of normality)
shapiro.test(alfai)
No hay evidencias de que la variable de efectos aleatorios manada no tenga distribución normal (p=...)
```

```
#3.2- homocedasticidad
# Con performance
check_homogeneity(m1)
```

```
## Grafico res vs predichos
plot(m1) # Res pearson vs valores predichos
```

Dado que la dispersión de los residuos parece constante y no se observan patrones en el gráfico de residuos vs. predichos, no hay evidencias para rechazar el supuesto de homocedasticidad.

```
#3.3- linealidad
#Aca no hay supuesto de linealidad porque son todas las predictoras cualitativas.
```

#Para cada conclusión indique de dónde obtiene la información (si es de un gráfico, de cuál, si es de una prueba analítica, de cuál, e informe p-valor si corresponde).

```
### Pregunta 4
#####
```

Analice los resultados del modelo ajustado e indique si las siguientes afirmaciones son V o F, justificando su decisión. Cuando corresponda, indique de qué salida obtuvo dicho resultado.

a) Los cambios en el HR entre sitios con y sin nieve difieren significativamente entre los 4 ambientes estudiados.

En el anova se obtiene una interaccion significativa, por lo que realizo un contraste simple viendo si hay diferencias significativas entre con y sin nieve, dentro de cada tipo de ambiente.

```
anova(m1)
library(emmeans)
(comp <- emmeans(m1, pairwise ~ nieve | areas))
```

F, los ambientes 1 2 y 4 muestran diferencias significativas entre las zonas "con" y "sin" nieve ($p < 0.0001$) pero el ambiente 3 no ($p = 0.7554$)

b) Si se ajustase un modelo marginal con matriz de correlación de simetría compuesta los coeficientes estimados por el summary() variarán ligeramente.

F, ajustar un modelo marginal con matriz de correlacion de simetria compuesta es equivalente a ajustar un modelo condicional en cuanto a los coeficientes estimados para variables con distribucion normal.

c) Las fuentes de variación aleatorias del modelo ajustado son dos: manadas e individuos.

#V, la varianza entre manadas de una misma combinación ambiente-nieve esta representada por σ^2_{manada} y la varianza entre individuos de una misma manada por σ^2 y la suma de ambas es la varianza aleatoria total.

d) La varianza del modelo es 0.88550 km².

F, ese valor corresponde al componente residual de la varianza del modelo, se obtiene del codigo summary(modelo). Las unidades son (km²)².
summary(m1)

e) Es indistinto evaluar la significación de los términos del modelo utilizando summary o anova.

F, en summary no se controla el error global y por lo tanto lo correcto es evaluar la significacion con el anova dado que hay una variable categorica con mas de 2 niveles.

Pregunta 5
#####

#Efectúe las pruebas que considere de mayor potencia para responder si se verifican las predicciones y sospechas de los investigadores.

#Los investigadores predicen que las zonas con nieve tendrán menor productividad de recursos alimenticios y por lo tanto menor cantidad de presas disponibles, por lo que los individuos de las manadas de esas zonas deberán tener un mayor HR.

#Sin embargo, sospechan que la disponibilidad de animales de cría puede atenuar estas diferencias.

```
options(emmeans = list(emmeans = list(infer = c(TRUE, TRUE)),contrast = list(infer =  
c(TRUE, TRUE))  
) # seteo de las opciones de salida  
# Primero se evaluamos si hay interaccion en el modelo.  
anova(m1)
```

Dado que la interaccion es significativa, la prueba de mayor potencia que permite responder las predicciones y sospechas de los investigadores es realizar comparaciones de efectos simples entre las dos condiciones de nieve para cada tipo de ambiente.

```
(comp <- emmeans(m1, pairwise ~ nieve | areas))
```

Las predicciones de los investigadores se verifican dado que para todos los ambientes el HR es significativamente menor ($p < 0.0001$, o límites del intervalo de confianza para las diferencias de media que no incluyen al 0) en zonas "sin nieve", a excepción del ambiente 3 que es el único con disponibilidad de animales de cría ("ganadería"), en donde no se encuentran diferencias entre los diferentes zonas de nieve ($p = 0.7554$).

Pregunta 6

#####

Informe e interprete el IC 95% para la magnitud de efecto del ambiente con el máximo cambio en el HR medio entre zonas con y sin nieve.

```
comp
```

Vemos que el ambiente con mayor estimate para la diferencia de condicion de nieve es el 4, con un valor de 29.991 km², y límites del intervalo de confianza [29.155 - 30.83] km².

Interpretación: Con una confianza del 95% se estima que en el ambiente 4 (pastizal) el home range esperado (o medio) es entre 29.155 y 30.83 km² mayor en zonas con nieve comparado con zonas sin nieve para lobos de una mandada típica.

Informe e interprete el HR medio estimado para el ambiente bosque nativo en condiciones de nieve.

```
#Bosque nativo = A1
```

```
comp$emmeans[1]
```

El home range medio estimado para un lobo de una manada típica en el tipo de ambiente bosque nativo en zonas con nieve es 40.1 km².

Pregunta 7

#####

Se sabe que el costo de identificar una manada en terreno es alto. ¿Haría alguna sugerencia orientada a aumentar los esfuerzos en incorporar más manadas o más individuos por manada en futuros estudios? Concluya en base a aportes a la variabilidad aleatoria del HR.

```
model_performance(m1, metrics = "ICC")
```

El valor del ICC es 0.013, esto indica que solo el 1.3% de la variabilidad aleatoria en el HR esta dada por la variacion entre manadas pertenecientes a una misma combinación ambiente-nieve, mientras que el 98,7% restante corresponde a la variacion aleatoria entre individuos de una misma manada.

Dado que el costo de identificar manadas es alto y la variabilidad entre manadas es baja respecto a la variabilidad entre individuos, sería más conveniente aumentar los esfuerzos en incorporar más individuos.

Pregunta 8

#####

Presente un gráfico con los principales resultados del ensayo con su epígrafe. Escriba un párrafo con la metodología estadística utilizada y las conclusiones del estudio.

```
library(ggeffects)
(a <- ggpredict(m1, terms = c("nieve", "areas")))
p <- plot(a, add.data = T, grid = F)
p + ggtitle("Home Range de lobos en manadas en diferentes ambientes y condiciones de
nieve") +
  labs(y = "HR(km2)", x = "Condicion de nieve")
```

Epígrafe. Figura 1. Home Range esperado (o medio) para manadas de lobos en diferentes tipos de ambiente y condición de nieve (media e IC 95%). A1: "bosque nativo", A2: "forestación", A3: "ganadería" y A4: "pastizal").

Metodología. Para la comparación del HR ajustó un modelo lineal general mixto con 2 variables predictoras fijas: tipo de ambiente (4 niveles, bosque nativo, forestación, ganadería y pastizal) y condición de nieve (2 niveles, con y sin nieve). Se incorporó la falta de independencia entre observaciones de una misma manada de forma implícita. El supuesto de normalidad de los errores y del efecto aleatorio se estudió mediante la prueba de Shapiro-Wilks y de forma gráfica (QQ plot). Para el supuesto de homocedasticidad se realizó un análisis gráfico de dispersión de residuos vs valores predichos. Se aplicaron 4 comparaciones a posteriori de efectos simples. Se trabajó con un $\alpha = 0,05$. Todos los análisis se efectuaron con R (R Core Team, 2023).

Conclusiones. Se obtuvo una interacción significativa entre condición de nieve y tipo de ambiente ($p < 0.05$), lo que significa que el cambio en el Home Range difiere entre zonas con y sin nieve según el ambiente en todos los ambientes. Para los ambientes bosque nativo, forestación y pastizal se encontró un aumento en el HR ($p < 0.0001$) en zonas con nieve comparado con zonas sin nieve. Sin embargo, en el ambiente con disponibilidad de animales de cría (ganadería) no se encontraron evidencias de cambio en el HR ($p = 0.7554$) entre las zonas con diferente condición de nieve. Estos resultados concuerdan con las predicciones de los investigadores, es decir que las manadas en zonas con nieve presentan mayor HR lo que puede ser explicado por la menor disponibilidad de alimento. Sin embargo, la disponibilidad de animales de cría puede explicar que no se detecten cambios en el HR en presencia de nieve. Es importante destacar que al tratarse de un estudio observacional no puede concluirse sobre causalidad y que la inferencia se limita a estos 4 ambientes estudiados.

Pregunta 9
#####

Completar el siguiente párrafo:
ranef(m1) # BLUP (efecto aleatorio) de cada camada

A partir de los datos recopilados en el estudio realizado se pueden obtener 24_ "ranef"
que corresponden a las variaciones que aporta cada manada.

El valor del CCI es de 0.013 y sus unidades son (no tiene unidades). Su interpretación, en contexto, es el que el 0.013 de la variabilidad aleatoria del modelo está dada por la variación entre manadas pertenecientes a una misma combinación ambiente-nieve. El valor máximo que puede tomar es 1 en cuyo caso significa que toda la varianza aleatoria estaría dada por la variación entre manadas y que dentro de cada manada los individuos serían homogéneos, es decir que tendrían la misma respuesta.

Las conclusiones de la parte fija del estudio son válidas para "todas las manadas de las 4 áreas de estudio". Esto es porque la inferencia de la parte fija se hace sobre los niveles estudiados de las variables fijas, que en este caso son 4 áreas distintas.

Pregunta 10
#####

Un investigador propone que los ejemplares más jóvenes tienden a tener HR más pequeños, por su temor a alejarse de la madriguera, y que esta relación es independiente del tipo de ambiente y de las condiciones climáticas. Al haber relevado el dato de la edad de cada individuo en campo, se decide incorporar como predictora a la edad del individuo (en meses).

¿Cómo se modifica el modelo planteado? Escriba el modelo en parámetros indicando si cambia la cantidad de parámetros estimada respecto al anterior.

Si incorpora coeficientes en el modelo, explíquelos en contexto.

Escriba la sintaxis en R del nuevo modelo (no tiene que ejecutarlo).

$HR_{ij} = \beta_0 + \beta_1 \times edad_{ij} + \beta_2 \times sin\ nieve_i + \beta_3 \times A2_i + \beta_4 \times A3_i + \beta_5 \times A4_i + \beta_6 \times sin\ nieve_i \times A2_i + \beta_7 \times sin\ nieve_i \times A3_i + \beta_8 \times sin\ nieve_i \times A4_i + M_i + Epsilon_{ij}$

i=1:24
j=1:4

β_0 : valor estimado de HR cuando edad=0, con nieve y en ambiente A1

β_1 : cambio medio en el HR por el aumento de un mes en la edad

β_2 : cambio esperado en HR respecto de β_0 cuando no hay nieve y el ambiente es A1

β_3 : cambio esperado en HR respecto de β_0 cuando hay nieve y el tipo de ambiente es A2

Beta4: cambio esperado en HR respecto de B0 cuando hay nieve y el tipo de ambiente es A3

Beta5: cambio esperado en HR respecto de B0 cuando hay nieve y el tipo de ambiente es A4

Beta6: cambio esperado en HR respecto de B0 cuando no hay nieve y el tipo de ambiente es A2

Beta7: cambio esperado en HR respecto de B0 cuando no hay nieve y el tipo de ambiente es A3

Beta8: cambio esperado en HR respecto de B0 cuando no hay nieve y el tipo de ambiente es A4

Este modelo requiere estimar 11 parametros, mientras que el anterior 10 parametros.

```
modelo <- lmer(HR ~ edad + nieve * areas + (1 | manada), datos)
```