

# Clase 14 Introducción Modelos Lineales Mixtos

Curso Introducción al Análisis de datos con R para la acuicultura.

Dr. José A. Gallardo y Dra. María Angélica Rueda.  
jose.gallardo@pucv.cl | Pontificia Universidad Católica de  
Valparaíso

31 July 2021

# PLAN DE LA CLASE

## **1.- Introducción**

- Repaso de modelos lineales generales.
- Modelos lineales mixtos (MLM) ¿Por qué y para qué?
- Efectos fijos y efectos aleatorios.
- Ecuación del modelo lineal mixto (MLM).
- Interpretación de MLM con R.

## **2.- Práctica con R y Rstudio cloud**

- Ajustar modelos lineales mixtos.
- Realizar gráficas avanzadas con ggplot2.
- Elaborar un reporte dinámico en formato pdf.

# MODELOS LINEALES GENERALES

Los modelos lineales generales son los modelos de regresión lineal simple, modelos de regresión lineal múltiple.

$$Y = X\beta + \epsilon$$

Efectos fijos ( $X\beta$ )

Efectos aleatorios ( $\epsilon$ )

Los modelos lineales generales deben cumplir los supuestos:

- Que no haya multicolinealidad (Regresión lineal múltiple).
- Que los residuos se distribuyan normal.
- Que haya homogeneidad de varianzas.

# MODELOS LINEALES MIXTOS

Los modelos lineales mixtos (MLM) son una generalización del modelo lineal de regresión clásico, contemplando la posible existencia de observaciones correlacionadas (ej. Medidas repetidas en el mismo individuo) o con variabilidad heterogénea, vinculadas a la presencia de factores aleatorios.

$$Y = X\beta + Zu + \epsilon$$

Efectos fijos ( $X\beta$ )

Efectos aleatorios ( $Zu + \epsilon$ )

Los modelos lineales mixtos surgen cuando no se cumplen los siguientes supuestos:

- Que hayan observaciones correlacionadas.
- Que NO haya homogeneidad de varianzas.

# ¿QUÉ SON EFECTOS FIJOS?

- ▶ Los efectos fijos se asumen que son determinados a propósito por el analista de los datos, eso dependerá de las variables a las que se les desea estimar efectos promedios.
- ▶ Los efectos fijos solo estiman medias de las variables predictoras.
- ▶ En un modelo lineal mixto las variables cuantitativas continuas o covariables pueden ser usadas como efectos fijos.

# ¿QUÉ SON EFECTOS ALEATORIOS?

- ▶ Los efectos aleatorios están asociados a grupos de observaciones. Los efectos aleatorios estiman varianzas.
- ▶ Para considerar una variable predictora cualitativa como un efecto aleatorio del modelo lineal mixto, dicha variable debe tener al menos 5 niveles.
- ▶ Una variable predictora categórica con dos niveles (binaria) NO puede ser un efecto aleatorio.
- ▶ Una variable aleatoria continua NO puede ser un efecto aleatorio.

# ALGUNOS EJEMPLOS DE EFECTOS ALEATORIOS

- i) Medidas repetidas sobre un mismo individuo (hay repeticiones).
- ii) Respuestas observadas en grupos de unidades experimentales homogéneas (bloques), pueden ser piscinas o estanques.
- iii) Mediciones de los animales (individuos) de una misma familia.

# ¿CÓMO SE PODRÍA DECIDIR SI ES EFECTO FIJO O ALEATORIO?

1). ¿Cuál es el número de niveles?

- ▶ Pequeño (Fijo)
- ▶ Grande o infinito (Posiblemente aleatorio)

2). ¿Son los niveles repetibles?

- ▶ Sí (Fijo)
- ▶ No (Aleatorio)

3). ¿Se necesitan realizar inferencias para niveles no incluidos en el muestreo?

- ▶ No (Posiblemente fijo)
- ▶ Sí (Posiblemente aleatorio)



# LIBRERÍA PARA AJUSTAR MODELOS LINEALES MIXTOS

```
library(lme4)
```

## Función **lmer()**

Dimensión de la base de datos: 578 observaciones y 5 columnas. Se evaluaron 4 dietas.

Table 1: Tabla de datos

ID	weight	time	animal	diet
1	42	0	1	1
2	51	2	1	1
3	59	4	1	1
4	64	6	1	1
5	76	8	1	1
6	93	10	1	1

# MODELO DE EFECTOS FIJOS

```
mod.1 <- lm(weight ~ diet + time, data = tilapia_diet)
```

Call:

```
lm(formula = weight ~ diet + time, data = tilapia_diet)
```

Residuals:

Min	1Q	Median	3Q	Max
-135.172	-17.154	-2.192	15.561	152.049

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.5428	4.0430	0.382	0.703
diet	11.7786	1.3054	9.023	<2e-16 ***
time	8.7653	0.2246	39.030	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 36.45 on 575 degrees of freedom

Multiple R-squared: 0.7379, Adjusted R-squared: 0.7369

F-statistic: 809.2 on 2 and 575 DF, p-value: < 2.2e-16

# MODELO DE EFECTOS FIJOS

```
mod.2 <- lm(weight ~ diet:time, data = tilapia_diet)
```

Call:

```
lm(formula = weight ~ diet:time, data = tilapia_diet)
```

Residuals:

Min	1Q	Median	3Q	Max
-99.16	-25.62	-14.72	13.69	191.18

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	59.74447	2.86909	20.82	<2e-16 ***
diet:time	2.57527	0.08941	28.80	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 45.54 on 576 degrees of freedom

Multiple R-squared: 0.5902, Adjusted R-squared: 0.5895

F-statistic: 829.6 on 1 and 576 DF, p-value: < 2.2e-16

# MODELOS LINEALES MIXTOS

```
mod.1r <- lmer(weight~ diet + time + (1|animal),  
               data = tilapia_diet)
```

```
Linear mixed model fit by REML ['lmerMod']  
Formula: weight ~ diet + time + (1 | animal)  
Data: tilapia_diet
```

```
REML criterion at convergence: 5601.8
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-3.0471	-0.5635	-0.1173	0.4864	3.4869

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
animal	(Intercept)	539.8	23.23
Residual		799.4	28.27

```
Number of obs: 578, groups: animal, 50
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	1.9520	7.7119	0.253
diet	11.7482	3.0039	3.911
time	8.7198	0.1754	49.701

```
Correlation of Fixed Effects:
```

	(Intr) diet
diet	-0.858
time	-0.229 -0.013

# MODELOS LINEALES MIXTOS

```
mod.2r <- lmer(weight~diet + time + (time|animal),  
               data = tilapia_diet)
```

```
Formula: weight ~ diet + time + (time | animal)  
Data: tilapia_diet
```

REML criterion at convergence: 4815.8

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-2.7463	-0.5697	-0.0478	0.5125	3.4703

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
animal	(Intercept)	168.31	12.973	
	time	14.13	3.758	-0.98
Residual		163.46	12.785	

Number of obs: 578, groups: animal, 50

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	23.018	2.668	8.626
diet	2.797	0.749	3.735
time	8.443	0.540	15.635

Correlation of Fixed Effects:

	(Intr) diet
diet	-0.620
time	-0.710 -0.005

# SELECCIÓN DE MODELOS DE EFECTOS FIJOS

Criterios de selección de modelos AIC y BIC usando anova()

```
anova(mod.1, mod.2)
```

```
Analysis of Variance Table
```

```
Model 1: weight ~ diet + time
```

```
Model 2: weight ~ diet:time
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	575	764036				
2	576	1194315	-1	-430280	323.82	< 2.2e-16 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# SELECCIÓN DE MODELOS MIXTOS

Criterios de selección de modelos AIC y BIC usando anova()

anova(mod.1r, mod.2r)

Data: tilapia\_diet

Models:

mod.1r: weight ~ diet + time + (1 | animal)

mod.2r: weight ~ diet + time + (time | animal)

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
mod.1r	5	5618.5	5640.3	-2804.3	5608.5			
mod.2r	7	4833.2	4863.7	-2409.6	4819.2	789.33	2	< 2.2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# RESUMEN DE LA CLASE

- 1). Revisión de conceptos: modelos lineales generales, modelos lineales mixtos.
- 2). Construir y ajustar modelos lineales mixtos.