

Clase 20 - Análisis de sobrevivencia

Diplomado en Análisis de datos con R para la Acuicultura

Dr. José Gallardo Matus

Pontificia Universidad Católica de Valparaíso

14 June 2022

PLAN DE LA CLASE

1.- Introducción

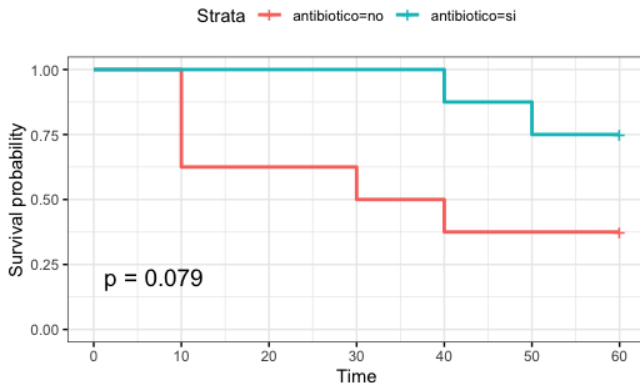
- ▶ Qué son los análisis de sobrevivencia?
- ▶ Método de Kaplan-Meier.
- ▶ Test estadístico.
- ▶ Estudios de caso
- ▶ Interpretación pruebas con R

2.- Práctica con R y Rstudio cloud.

- ▶ Realizar análisis de sobrevivencia.
- ▶ Realizar gráficas avanzadas con ggplot2.
- ▶ Elaborar un reporte dinámico en formato PPT.

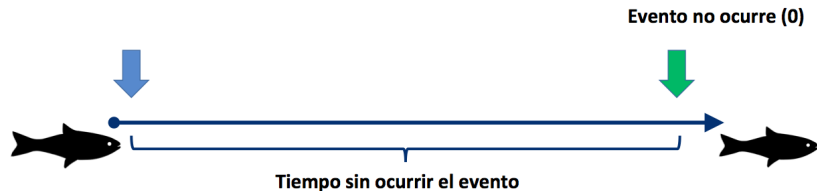
ANÁLISIS DE SUPERVIVENCIA

Conjunto de herramientas estadísticas No paramétricas utilizadas para analizar la probabilidad de que un evento (muerte/falla) ocurra en un determinado Tiempo.



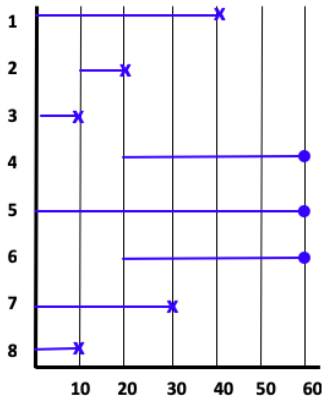
CONCEPTOS RELEVANTES

Variables respuesta: Tiempo de supervivencia y Estado (0 - 1).



CONCEPTOS RELEVANTES: DATOS CENSURADOS

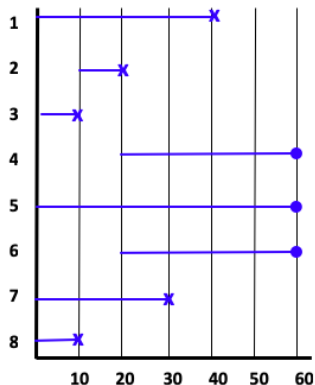
- ▶ Los organismos pueden entrar en diferentes tiempos al estudio.
- ▶ El evento puede ocurrir después de finalizar el estudio (Censurar datos).



| Ind. | Tiempo | Estado |
|------|--------|--------|
| 1 | 40 | 1 |
| 2 | | |
| 3 | | |
| 4 | 40 | 0 |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |

X = evento
● = Censura

EJEMPLO SET DE DATOS



| Ind. | Tiempo | Estado |
|------|--------|--------|
| 1 | 40 | 1 |
| 2 | 10 | 1 |
| 3 | 10 | 1 |
| 4 | 40 | 0 |
| 5 | 60 | 0 |
| 6 | 40 | 0 |
| 7 | 30 | 1 |
| 8 | 10 | 1 |

X = evento

● = Censura

MÉTODO DE KAPLAN - MEIER

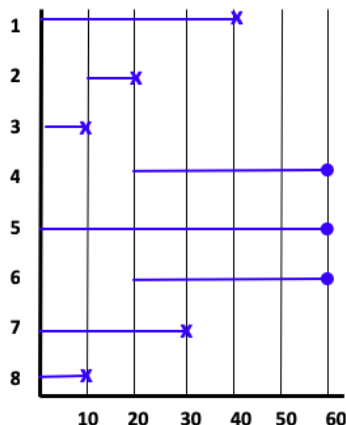
El método de Kaplan-Meier es un método no paramétrico que estima las probabilidades de supervivencia $S(t)$ en los instantes en los que ha ocurrido el evento.

$$\hat{S}(t) = \prod_{t < t_1} \frac{n_i - d_i}{n_i}$$

d_i , el número de muertes en el momento t_i

n_i , el número de sujetos en riesgo justo antes de t_i .

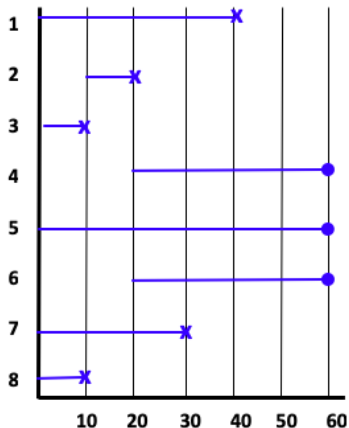
CALCULAR PROBABILIDAD KM



| Tiempo | Prob. sobrevivir | Estimador K-M |
|--------|------------------|---------------|
| 0 | $5/5 = 1$ | 1 |
| 10 | $4/6 = 0,66$ | 0,66 |
| 20 | $5/6 = 0,83$ | 0,54 |
| 30 | $4/5 = 0,80$ | 0,43 |
| 40 | | |
| 50 | | |
| 60 | | |

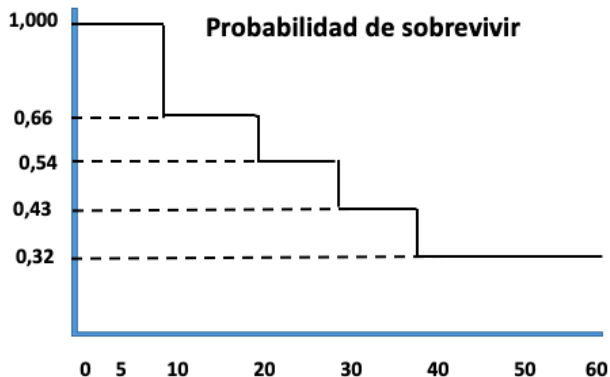
$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i}$$

CALCULAR PROBABILIDAD KM



| Tiempo | Prob. sobrevivir | Estimador K-M |
|--------|---------------------|------------------|
| 0 | $5/5 = 1$ | 1,00 |
| 10 | $4/6 = 0,66$ | 0,66 |
| 20 | $5/6 = 0,83$ | 0,54 |
| 30 | $4/5 = 0,80$ | 0,43 |
| 40 | $3/4 = 0,75$ | 0,32 |
| 50 | $3/3 = 1$ | 0,32 |
| 60 | $3/3 = 1$ | 0,32 |

GRÁFICA DE SOBREVIVENCIA.



| T | K-M |
|----|------|
| 0 | 1,00 |
| 10 | 0,66 |
| 20 | 0,54 |
| 30 | 0,43 |
| 40 | 0,32 |
| 50 | 0,32 |
| 60 | 0,32 |

PRUEBA ESTADÍSTICA PARA COMPARAR TRATAMIENTOS

Test estadístico no paramétrico Log rank test.

$$G = 2 \sum_i O_i \ln \frac{O_i}{E_i}$$

Hipótesis

H₀: No existen diferencias entre los grupos en la probabilidad de que ocurra un evento (muerte) en ningún tiempo.

H₁: Existen diferencias entre los grupos en la probabilidad de que ocurra un evento (muerte) en algún tiempo.

| | Grupo 1 | Grupo 2 |
|---------|------------|------------|
| Evento | Prob. KM A | Prob. KM C |
| Sensura | Prob. KM B | Prob. KM D |

Fuente: The log rank test

PRÁCTICA ANÁLISIS DE DATOS

- ▶ El trabajo práctico se realiza en Rstudio.cloud.

RESUMEN DE LA CLASE

- ▶ Revisión de análisis de supervivencia y tiempos de vida media o falla.
- ▶ Cálculo de probabilidad mediante método de Kaplan- Meier.
- ▶ Ejecutamos e interpretamos resultados test de supervivencia con R.