

# Defunciones de menores de una semana por causa y sexo

Luis Miguel Calvo Magaz y Adrián Poncela Gómez

24 de Abril de 2019

## 1 Introducción: Análisis del conjunto de datos

Los datos utilizados para la realización de este trabajo de Análisis de supervivencia, se trata de un conjunto de datos obtenidos del INE. Apartado: Sociedad-Salud-Defunciones según la Causa de Muerte 2017.  
**Defunciones de menores de una semana por causas, sexo y edad.**

Esos datos muestran, el número de fallecidos por cada día, dentro de la primera semana de vida de los bebés. Hay que tener en cuenta, que todos los individuos que se estudian, fallecen en este estudio. No se estudia ningún individuo que viva más de una semana. Los datos se muestran divididos por sexo y edad de fallecimiento.

La edad de fallecimiento está calculada por la diferencia entre la fecha de defunción y la de nacimiento, a excepción de los menores de 24 horas de vida. Por tanto, como fallecidos de un día figuran sólo aquellos que han vivido más de 24 horas y fallecido al día siguiente.

**Variable objeto de estudio:** Edad (días) de fallecimiento de los bebés.

**Origen:** 1 de Enero de 2017

Suceso de interés: Fallecimiento del bebé.

**Escala del tiempo:** Desde el 1 de Enero de 2017 - 31 de Diciembre de 2017.

**Censuras y truncamientos:** Para este conjunto de datos, no tenemos censuras ni truncamientos ya que todos los individuos que son objeto de estudio, mueren, ninguno sobrevive.

Dentro de este conjunto de datos, además del sexo, se nos proporciona otra covariable más, la causa de muerte.

Las causas de muerte aparecen divididas en 4 grupos, y esos grupos a su vez en subgrupos, hemos decidido considerar como covariable esos 4 grupos principales y no sus subdivisiones ya que considerar más de 90 casos de muerte para tan pocos datos (490 en un año), no lo consideramos relevante.

Los apartados de causas de muerte considerados, siguiendo la división que utiliza el INE, son los siguientes:

**Perinatal** (383 individuos)

**Malformaciones** (86 individuos)

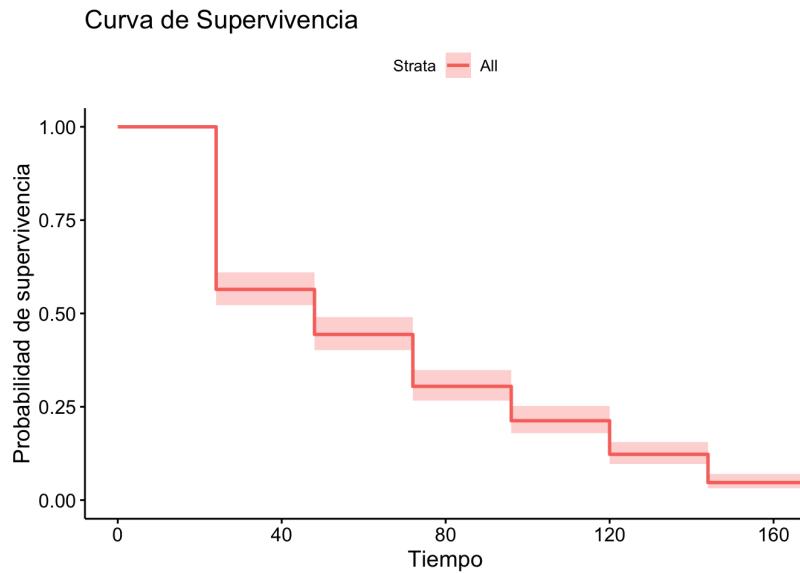
**Enfermedades recién nacido** (20 individuos)

**Causas externas** (1 individuo)

Sin embargo, sólo un bebé se considera dentro del apartado de "Causas externas". Visto que no hay una similitud, de este individuo con las demás causas, y al ser sólo una observación, hemos decidido prescindir de él para la realización del estudio.

Respecto a los 3 tipos de causa que nos quedan, vemos como la distribución de los individuos en cada una de las causas, es igual, ya que una causas tiene más individuos que las demás. Se ve que hay una gran diferencia entre ellas. Hay 4 veces más individuos muertos por malformaciones que por enfermedades de recién nacido. Además Hay casi 5 veces más muertos por perinatal que por malformaciones.

## 2 Curva de supervivencia sin covariables



En esta curva de supervivencia estimada con el método de Kaplan-Meier podemos ver cómo en las primeras 24 horas se alcanza el mayor número de defunciones de recién nacidos. A partir de ahí, el número de muertes disminuye, decreciendo linealmente. Los intervalos de confianza se estrechan a medida que avanza el tiempo. El número de muertes, por cada día, disminuye con el paso del tiempo. La curva finaliza con valor cero, ya que todos los recién nacidos de este estudio, han fallecido.

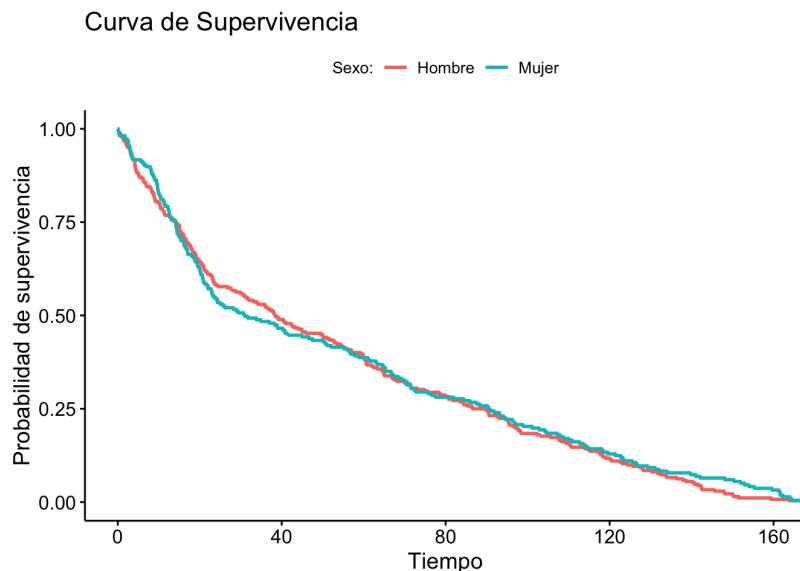
Tras observar esta curva descartamos completamente el usar un modelo de riesgos proporcionales de COX, puesto que a simple vista se puede ver cómo los riesgos no son proporcionales.

## 3 Distribución aleatoria uniforme de la variable objeto del estudio

Al tener sólo los datos para valores puntuales (Muertes entre las 0 y las 24 horas, entre las 24 y las 48, etc), se ha decidido realizar una simulación utilizando la distribución uniforme para cada rango.

La idea es distribuir las muertes a lo largo del día como si fuese la edad real de muerte y así simular una curva más realista a la que sería en la realidad.

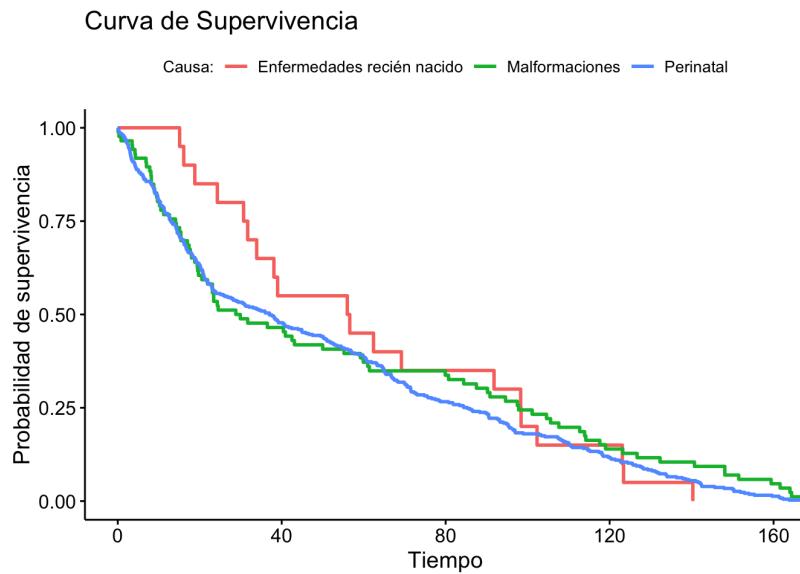
### 3.1 Curva de supervivencia con la covariable sexo



Los resultados con la simulación de la uniforme son diferentes en cuanto a las gráfica anterior, en primer lugar, la distribución ya no es escalonada. Se muestra una función más lineal y por lo tanto más real, aunque manteniendo las interpretaciones que concluimos anteriormente.

Apenas hay diferencias entre hombres y mujeres, esto es un dato que llama bastante la atención, ya que para un número de observaciones tan bajo (490) podría existir una diferencia más considerable en cuanto al número de fallecidos en la primera semana.

### 3.2 Curva de supervivencia con la covariable causa de fallecimiento



En este gráfico, podemos ver dos partes. Una que es antes de las 72 horas, que serían los 3 primeros días. Y de los 3 días en adelante, es decir, más de 72 horas. En la primera parte del gráfico se ve como hay dos curvas muy similares, que son las de Malformaciones y Perinatal, que están muy juntas. Sin embargo, la tercera causa que se corresponde con las enfermedades del recién nacido, va por encima.

Se considera conveniente un estudio de esas tres primeras horas por separado y comprobar si esas dos covariables (Malformaciones y perinatal) son realmente tan similares como para considerarlas iguales.

### 4 Test log-rank para Malformaciones y Perinatal

Se decide aplicar un test no paramétrico log-rank para comparar las dos causas tan similares, causa perinatal y malformaciones.

```
n=469, 1 observation deleted due to missingness.
```

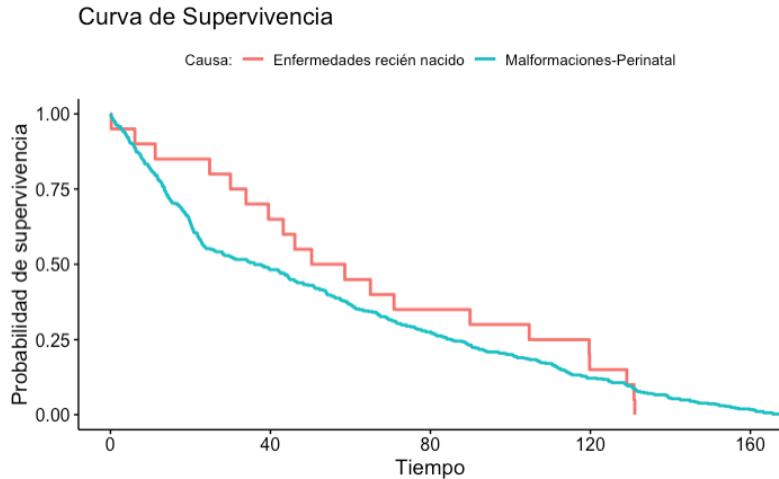
	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
causa=Malformaciones	86	86	92.4	0.446	0.565
causa=Perinatal	383	383	376.6	0.110	0.565

```
ChiSq= 0.6 on 1 degrees of freedom, p= 0.5
```

No se puede rechazar que ambas causas sean iguales por lo que se procede a juntarlas en una sola y comparar únicamente dos grupos de causas

### 5 Comparación de dos causas principales

Como se ha dicho en la sección anterior, las causas perinatal y malformaciones se pueden juntar en una única causa y comparar esta con "Causas externas"



A simple vista se puede observar una gran diferencia entre las curvas de supervivencia entre ambas causas, sin embargo al realizar un test no paramétrico log-rank para comparar las dos causas los resultados son los siguientes:

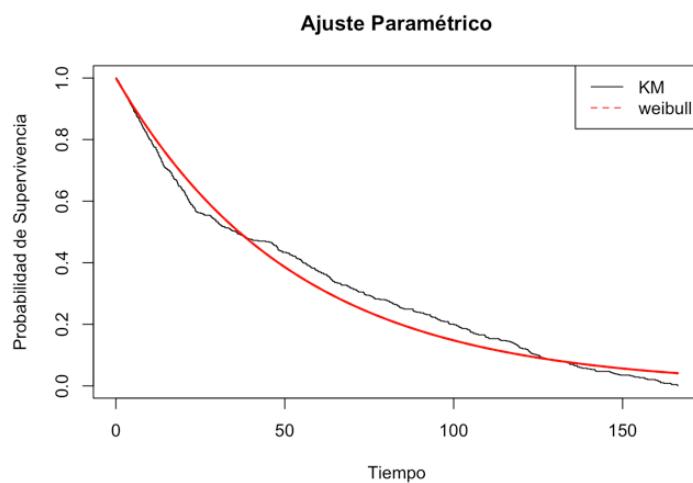
	N	Observed	Expected	$(O-E)^2/E$	$(O-E)^2/V$
causa=Enfermedades recién nacido	20	20	23	0.385	0.407
causa=Perinatal-Malformaciones	469	469	466	0.019	0.407

Chisq= 0.4 on 1 degrees of freedom, p= 0.5

No se puede rechazar la hipótesis de que ambos grupos son iguales, esto puede ser debido a la diferencia tan grande en el número de observaciones.

## 6 Estimación de la curva de supervivencia por una Weibull

Los datos aproximados uniformemente tiene una forma parecida a la de una distribución Weibull salvo porque tiene un decrecimiento un poco más lineal en etapas medias del estudio. La Weibull que mejor se ajusta a los datos sería una Weibull de parámetros: shape = 1.0085 y scale 53.1075.



## **7 Puntos a destacar sobre el estudio**

Se observa una gran mortalidad los dos primeros días y luego se suaviza.

Resulta muy interesante no encontrar a penas diferencia en la mortalidad entre sexo masculino y femenino. Diferencias visuales entre la causa de enfermedades de recién nacido y Perinatal-Malformaciones, sin embargo no se pueden demostrar estadísticamente.

Las diferencias entre causas son más destacables las primeras 72 horas, a partir del cuarto a penas se pueden distinguir.

La distribución de Weibull que mejor ajusta el modelo sería una Weibull(1, 53).