

# Prácticas de Fiabilidad

## Práctica 3:

### **Objetivo:**

En esta práctica se van a ajustar datos censurados.

Vamos a calcular el estimador de Kaplan-Meier y también modelos Weibull de varias variables.

Una vez que hayamos ajustado los datos, podremos saber la duración media de un componente y conocer tanto su tasa de fallos como su función de supervivencia.

Conocida la función de supervivencia podremos evaluar la probabilidad de que el componente dure más de un determinado tiempo.

### **Datos:**

Los datos que se van a analizar se encuentran en el fichero *practica 3 fiabilidad.sf*. El fichero contiene datos las variables:

1. Tipo: Tipo de tratamiento recibido.
2. Tiempos censurados: La duración del proceso.
3. Censura: La variable toma el valor 1 si el tiempo es censurado. En otro caso toma el valor cero.

### **Qué hay que hacer:**

#### **1. Estimador de Kaplan Meier**

- Se abre el fichero *practica 3 fiabilidad.sf*.
- Se va a:

**DESCRIBE**

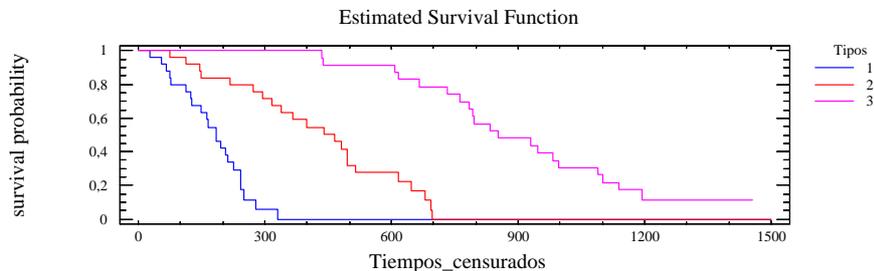
**Life data**

**Life tables (times)**

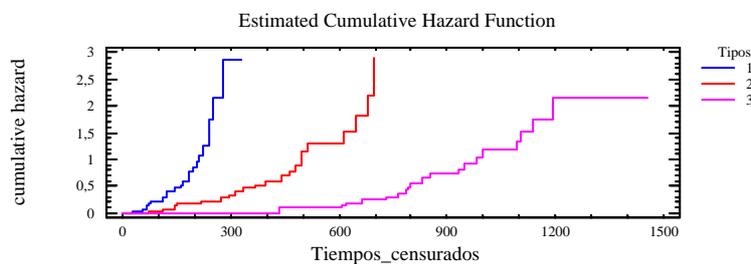
- En Data ponemos el nombre de la variable que queremos analizar: tiempos.
- En Censored se pone la variable que codifica las censuras (La variable que toma valores cero o uno).
- En Group se pone la variable que codifica los grupos (en caso de que haya varios)
- El ordenador proporciona una serie de gráficos que van a ser de ayuda en la toma de decisiones:

## Estimated Survival Function Tasa de fallos acumulada

El primero (**Estimated Survival Function**) proporciona la supervivencia estimada mediante el estimador de Kaplan Meier.



El segundo (**Tasa de fallos acumulada**) sirve para detectar si el componente tiene tasa de fallos creciente, decreciente o constante.



En este caso y tal como se vio en la práctica 1 las tasas de fallos parecen crecientes.

Además el ordenador proporciona los valores del estimador de Kaplan Meier para cada grupo y la duración media estimada.

En el ejemplo que se está estudiando, la media es:

Grupo 1: Duración media = 177,05  
Grupo 2: Duración media = 426.16  
Grupo 3: Duración media = 914.2

Si no se hubiera tenido en cuenta la censura, y se hubiesen calculado las duraciones medias utilizando los valores de tiempos observados, estas duraciones hubieran sido:

Grupo 1: Duración media sin censura = 169  
Grupo 2: Duración media sin censura = 406  
Grupo 3: Duración media sin censura = 853

Como puede observarse, se subestima la duración de los componentes.

Pinchando el icono TABULAR OPTIONS, se puede elegir la opción GROUP COMPARISONS.

Aquí obtenemos dos tests estadísticos (logrank y Wilcoxon) que contrastan si los grupos son significativamente distintos o no.

El resultado se interpreta a partir del p-valor. Cuanto más pequeño sea, más evidencia existe de diferencias entre los grupos. Si el p-valor es menor de 0.05 pensaremos que existen diferencias entre los grupos.

En el caso estudiado hay fuerte evidencia de que hay diferencias por grupos.

## 2. Ajuste paramétrico del modelo

Vamos a

### Describe Life Data Weibull análisis

Introducimos los datos de la misma manera que en el análisis anterior. Obtenemos:

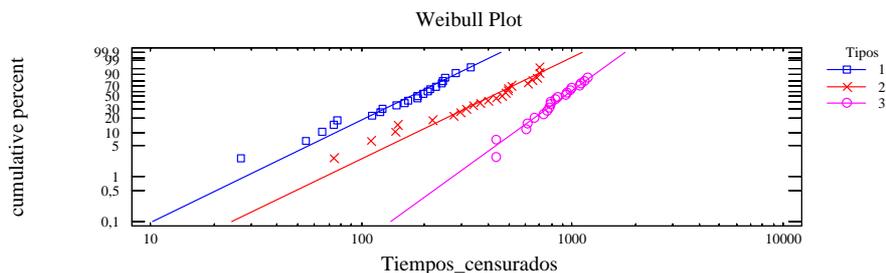
Analysis Summary

```
Data variable: Tiempos_censurados
Censoring: Censura
Groups: Tipos
Number of groups = 3
Estimation method: maximum likelihood
```

Group	Sample Size	Number of Failures	Estimated Shape	Estimated Scale	Starting Point
1	25	22	2,32743	200,625	0,0
2	25	22	2,31106	482,54	0,0
3	25	20	3,46217	1015,8	0,0

Donde Estimated shape proporciona las  $\beta$  para los grupos que son 2.32 2.31 y 3.46 respectivamente. Los parámetros de escala  $\lambda$  son 200,62 482,5 y 1015,8 respectivamente.

También se obtiene el gráfico de datos Weibull. La diferencia respecto a la práctica anteriores es que ahora hay datos incompletos.



En nuestro ejemplo los datos aparecen alineados, por lo que el modelo Weibull es adecuado.

Para calcular las probabilidades de tener que esperar más un tiempo  $t$ , hay que pinchar el icono amarillo (Tabular options) e ir a Tail Areas.

Tail areas proporciona:

$\Pr(T \leq t)$  es decir  $1-S(t)$ .

La probabilidad de durar más de 500 será:

$$\Pr(T \leq 500) = 0.99 \text{ para el grupo 1}$$

$$\Pr(T \leq 500) = 0.66 \text{ para el grupo 2}$$

$$\Pr(T \leq 500) = 0.08 \text{ para el grupo 3}$$

por tanto la probabilidad pedida,

$$S(500) = \Pr(T \geq 500) = 1 - 0.99 = 0.01 \text{ para el grupo 1}$$

$$S(500) = \Pr(T \geq 500) = 1 - 0.66 = 0.34 \text{ para el grupo 2}$$

$$S(500) = \Pr(T \geq 500) = 1 - 0.08 = 0.921 \text{ para el grupo 3}$$

### **Ejercicio a Entregar:**

Ajustar un estimador Kaplan Meier y un modelo Weibull para los datos de Edad de jubilación de una muestra de trabajadores obtenida de la EPA de 1987. El grupo va a ser si el trabajador estaba parado antes de jubilarse (PARADO=1) o estaba trabajando (PARADO=0)

La variable CENSJUBILADOS=1 representa un dato censurado.

Dibujar esquemáticamente su estimador K-M. Calcular la media para ambos grupos. ¿Cuál es la probabilidad de que un trabajador se jubile antes de los 65 años estando parado? ¿Y si trabajaba?

Realizar el análisis Weibull. ¿Cuál es la probabilidad de que un trabajador se jubile antes de los 65 años estando parado? ¿Y si trabajaba? ¿Tiene algún problema el análisis weibull?