



UNIVERSIDAD DE GRANADA

TRABAJO FIN DE MÁSTER

GRADO EN ESTADÍSTICA

EL PROCESO DE POISSON EN FIABILIDAD

Autora:

Sinia Arely Galvez Alvarado

Tutoras:

María Luz Gámiz Pérez

Rocío Raya Miranda



Facultad de
Ciencias

Febrero 2022

Tabla de contenidos

1. El proceso de Poisson	5
1.1. El proceso de Poisson homogéneo	5
1.1.1. Propiedades básicas del proceso de Poisson homogéneo	9
1.1.2. Distribución de los tiempos entre llegadas y de los tiempos de llegadas	10
1.1.3. Sucesión de tiempos de espera	11
1.1.4. Distribución condicional de los tiempos de llegadas . .	11
1.1.5. Tiempos residuales	12
1.2. El proceso de Poisson no homogéneo	13
1.3. Algunas propiedades del PPNH	16
1.3.1. Truncamiento en número de fallos	16
1.3.2. Truncamiento en tiempo de fallos	18
1.4. El PPNH en Fiabilidad. La función ROCOF	19
1.4.1. El concepto de reparación mínima	19
1.4.2. El modelo: Proceso de Poisson no homogéneo (PPNH)	22
2. Inferencia paramétrica en el PPNH	26
2.1. Datos de recurrencia en fiabilidad	26
2.2. Estimación máximo verosimil en PPNH	28
2.2.1. Función de verosimilitud en el proceso de Poisson . .	28
2.2.2. Función de verosimilitud del proceso de Poisson superpuesto	29
2.3. Inferencia en el PP con tasa de fallo potencial	29
2.3.1. Estimación puntual de los parámetros: β y θ	30
2.3.2. Estimación por intervalos y contrastes de hipótesis . .	32
2.4. Estimación ML para el PPNH con tasa de fallo log-lineal . . .	36
2.5. Tratamiento computacional: implementación en R	37
2.6. Pruebas de la tendencia de la tasa de fallo	39
2.6.1. El test <i>Military Handbook</i>	39
2.6.2. El test de <i>Laplace</i>	40
2.6.3. El test de Lewis-Robinson	40

2.7.	Prueba de tiempos de interurrencia independientes	41
2.8.	Generación de trayectorias muestrales de un PPNH	41
2.8.1.	PPNH con una tasa de fallo potencial	42
2.8.2.	PPNH con una tasa de fallo log-lineal	42
2.8.3.	Generación de trayectorias muestrales de procesos de Poisson en R	43
2.8.4.	Ejemplo práctico	46
3.	Inferencia no paramétrica en el PPNH	49
3.1.	Estimación suave de la función ROCOF	49
3.1.1.	Descripción del modelo	49
3.1.2.	Implementación en R	51
3.2.	El método de validación cruzada	52
3.2.1.	Descripción del modelo	52
3.2.2.	Implementación en R	54
3.2.3.	Ejemplo práctico	55
3.3.	Bandas de confianza para la función ROCOF	56
3.3.1.	Descripción del modelo	56
3.3.2.	Método de remuestreo	57
3.3.3.	Aproximación Bootstrap	58
3.3.4.	Bandas de confianza	58
3.3.5.	Implementación en R	59
3.4.	Estimación con múltiples realizaciones	61
3.4.1.	Descripción del modelo	61
3.4.2.	Implementación en R	63
3.4.3.	Ejemplo práctico	67
4.	Aplicación Práctica	69
4.1.	Descripción general de los datos	69
4.2.	Ajuste paramétrico	71
4.3.	Ajuste no paramétrico	74
4.3.1.	Análisis de trayectorias individuales	74
4.3.2.	Estimación de la tendencia con multiples realizaciones	76

Resumen

Un proceso de Poisson homogéneo es un proceso estocástico de recuento, en donde las ocurrencias de los sucesos están determinadas por tres hipótesis: homogeneidad que nos indica que k llegadas en un intervalo $(t, t + h]$ depende, solamente, de la variable k y h , mas no de t ; sin memoria, es decir, el proceso de Poisson es con incrementos independientes; y, regularidad, ya que no hay llegadas simultáneas. Las tres propiedades no se presentan en distintos problemas de la vida real, pero, son una excelente opción para empezar con el estudio de esos problemas. $\{N(t), t \geq 0\}$ es un proceso de Poisson, de modo que $N(t)$ representa el número de ocurrencias de un suceso en el intervalo $[0, t]$, la variable aleatoria $N(t)$ se distribuye según una ley de Poisson de parámetro λt , se tiene que $E[N(t)] = \lambda t$ de ahí que λ se denomine la razón o intensidad del proceso, y puede interpretarse como el número medio de llegadas por unidad de tiempo. En un PP, $\{T_n; n \geq 1\}$ denota los tiempos entre llegadas del proceso y, T_n con $n = 1, 2, \dots$ son independientes e idénticamente distribuidos con distribución exponencial de media $\frac{1}{\lambda}$, además, se denota S_n al tiempo en que ocurre la n -ésima llegada del proceso, $\{S_n, n \geq 1\}$ es la sucesión de tiempos de espera del proceso. La hipótesis de incrementos independientes estacionarios es equivalente a suponer que en cualquier instante de tiempo el proceso se reinicia probabilísticamente: a partir de cualquier punto, el proceso es independiente de todo lo que ha ocurrido previamente y tiene la misma distribución que el proceso original. Es decir, el proceso tiene la propiedad de no memoria, y por tanto es esperable que los tiempos entre llegadas tengan distribución exponencial. Si durante un tiempo t ha ocurrido una llegada según el proceso de Poisson de parámetro λ , y se desea calcular la distribución del instante de la llegada, debido a que es un proceso con incrementos estacionarios e independientes, cada subintervalo en $[0, t]$ de la misma amplitud tiene la misma probabilidad de contener tal llegada y la variable T_1 condicionada al suceso $N(t) = 1$, sigue una distribución uniforme en $[0, t]$.

En el Proceso de Poisson no Homogéneo en sus siglas PPNH se habla de llegadas aleatorias, que no son a tasa constante en el tiempo, es decir, la función de intensidad es variable en el tiempo t . Cuando un proceso de conteo se realiza con incrementos independientes y no estacionarios se puede ver como un PPNH. Como no se requieren incrementos estacionarios, se acoge la posibilidad de que las llegadas se produzcan más probablemente en ciertos intervalos de tiempo que en otros. Por definición, $\Lambda(t)$, la función de intensidad acumulada, se define como $\Lambda(t) = \int_0^t \lambda(s) ds$, una función continua y no decreciente. Este último teorema nos establece que todo PPNH puede asociarse a un PPH. Los PPNHs son buenos modelos para sistemas sujetos

a fallos y reparaciones. Al tratarse de sistemas reparables (Reparación Mínima), consideramos que el sistema sufre múltiples fallos a lo largo del tiempo, por lo que necesitamos un modelo estocástico que describa la ocurrencia de eventos en el tiempo, es decir, un proceso puntual. En general, el tiempo transcurrido entre los sucesivos fallos no es independiente ni está idénticamente distribuido, como ocurre con los modelos de Reparación Perfecta, de los que un caso particular es el PPH. La distribución de un PPNH y, en consecuencia, el comportamiento de la fiabilidad a lo largo del tiempo de un sistema mínimamente reparado, se caracteriza completamente por la función ROCOF. La forma de la ROCOF puede mostrar muchos tipos de comportamiento y las posibilidades son innumerables. En este sentido, es interesante el trabajo de Krivtsov (2007), que considera alternativas a los dos modelos tradicionalmente y utilizados y que se denominan modelo *log-lineal* y modelo *power law*; propone el uso de las formas paramétricas correspondientes a un gran número de familias de distribución tradicionales para la función de distribución subyacente del proceso. Sin embargo, los modelos paramétricos no suelen ser lo suficientemente flexibles como para cubrir todas las formas posibles de la función de intensidad y, por tanto, los procedimientos no paramétricos resultan muy atractivos.

El modelo que describe una población de sistemas, está basado en la función acumulativa media *MCF* a la edad t del sistema; *MCF* es definida como $\Lambda(t) = E[N(t)]$, donde la esperanza es sobre la variabilidad de cada sistema y la variabilidad unidad en la población. El proceso de ley de potencia es un PPNH con función de intensidad que tiene la forma $\lambda(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}$. El parámetro β afecta a la forma en que el sistema se deteriora o mejora con el tiempo. La no linealidad de un gráfico de datos de recurrencia de un sistema indica que los tiempos de interrecurrencia no están idénticamente distribuidos. Un gráfico de tiempos de interrecurrencia frente a la edad del sistema o un "gráfico de series temporales" que muestre los tiempos de interrecurrencia frente al número de recurrencias permitirá descubrir tendencias o ciclos que sugieran que los tiempos de interrecurrencia no están distribuidos de forma idéntica. Existen varias pruebas formales de tendencia de la tasa de fallo. Cuando se elige un modelo de ley de potencia del PPNH o sin tendencia, la prueba de *Military Handbook* es mejor para encontrar significancia. La prueba de Laplace es mejor para encontrar significancia cuando la elección es entre sin tendencia y un modelo exponencial PPNH. Tanto la prueba del *Military Handbook* como la prueba de *Laplace* pueden dar lugar a conclusiones erróneas en situaciones en las que no hay tendencia. Las pruebas de Laplace y Lewis-Robinson son más eficientes cuando las muestras grandes siguen aproximadamente una distribución *NOR*(0, 1). Para lograr la estima-

ción no paramétrica de la ROCOF se considera una o varias realizaciones de un PPNH, se observa un único sistema o, una población de sistemas con las mismas características. No se adopta ningún supuesto para la forma funcional de la función ROCOF $\lambda(t)$, excepto que es una función suave, definiéndose la suavidad en términos de algunas propiedades de derivabilidad de la función con respecto a t . Para estimar el parámetro de suavizado, se considera el método de validación cruzada por mínimos cuadrados que es un procedimiento completamente automático (o basado en datos) para elegir un valor adecuado para el parámetro de ancho de banda.

En nuestro estudio tomamos un conjunto de 68 sistemas con una amplitud constante en una aleación de aluminio para investigar la propagación de grietas por fatiga. Se toma al vector tiempo t como variable independiente y, esta se mide en ciclos de funcionamiento; la longitud de la grieta será la variable dependiente. Modelizamos el fenómeno usando un proceso de Poisson $\{N(t)\}$, donde $N(t)$ tiene un salto de tamaño 1 y cada vez que la grieta se incrementa una cantidad fija de $0,2mm$, siendo este el suceso de interés, entendemos que la grieta aumenta de modo continuo, pero, definimos el proceso como si el sistema recibiese golpes en momentos determinados que incrementan la grieta en esa magnitud. Una vez graficadas las trayectorias de los sistemas, observamos que los gráficos llevan una forma exponencial y utilizando las pruebas de tendencia Laplace y MH para dos sistemas, obtuvimos como resultado que no había tendencia constante. Seguidamente se hizo un ajuste paramétrico con los modelos potencial y exponencial a dos de las muestras; determinando que el ajuste potencial resulta ser más significativo, en comparación con el exponencial. La tendencia de los gráficos alertó sobre la no linealidad del modelo. Se realizó el análisis no paramétrico, se crean las curvas para los sistemas que representan las funciones score correspondiente al método de validación cruzada. Se alcanza un mínimo global, lo que determina el valor del ancho de banda óptimo según este criterio que proporcionará la curva que estima la función ROCOF minimizando el (ISE). Se representan las ROCOF para cada muestra y un promedio de todas las curvas. Se vio que desde el principio esta función crece de modo que los fallos van produciéndose con más asiduidad conforme van acumulándose ciclos de funcionamiento. En los sistemas se alcanza un máximo alrededor de los 200000-250000 ciclos, a partir de ese momento la curva decrece rápidamente. Podríamos pensar en un modelo paramétrico que sugiera esta curva, considerando al modelo Log-Normal. Se siguió el procedimiento de Bartoszyński, para aproximar el valor de la función λ . La función ROCOF acumulada es suave, continua y creciente a lo largo de todo t . Los sistemas se van deteriorando a medida pasa el tiempo, algo que nos indicaban los datos, ya que cada vez que había un fallo, había un incremento en el tamaño de la grieta

y parece apreciarse un cambio de tendencia en la tasa de fallos acumulada que alrededor del mismo punto donde aparece el pico en la tasa ordinaria la función parece aminorar el ritmo de crecimiento.

Capítulo 1

El proceso de Poisson

1.1. El proceso de Poisson homogéneo

Un proceso de Poisson homogéneo es un proceso estocástico de recuento, en donde las ocurrencias de los sucesos están determinadas por tres hipótesis:

- (i) Homogeneidad: la probabilidad de que ocurran k_1, k_2, \dots, k_n llegadas en n cualesquiera intervalos de tiempo disjuntos, depende sólo de k_i y del tamaño de esos intervalos, no de la posición que ocupan en la recta de tiempos. De manera específica, la probabilidad de que ocurran k llegadas en un intervalo $(t, t + h]$ depende, solamente, de la variable k y h , mas no de t . En este sentido, se dice que el proceso homogéneo tiene incrementos estacionarios.
- (ii) Sin memoria: la probabilidad de que ocurran k llegadas en el intervalo $(t, t + h]$ es independiente del número de llegadas que se han realizado hasta t . Es decir que, el proceso de Poisson es un proceso con incrementos independientes. Las variables número de ocurrencias en intervalos de tiempo disjuntos son independientes en probabilidad.
- (iii) Sin llegadas simultáneas: el número de llegadas en un intervalo de longitud infinitesimal tiende en probabilidad a 0. Por tanto, las trayectorias muestrales son aplicaciones con discontinuidades de salto, estos últimos se producen en los instantes de llegada y tienen magnitud uno. No hay posibilidad de que ocurran dos o más llegadas en el mismo instante de tiempo.

Resumiendo las tres hipótesis anteriores, se podría decir que un proceso de Poisson es un proceso de valores enteros con incrementos independientes estacionarios que tiene saltos unidad.

En la vida real, frecuentemente estas tres propiedades no necesariamente se cumplen. El número de clientes que llega a un banco no es homogéneo durante el intervalo de apertura de ese lugar. Aunque, sobre pequeños intervalos de tiempo, la homogeneidad sería una buena aproximación. En un call center, una llamada telefónica puede afectar el próximo conjunto de llamadas entrantes, por lo que la propiedad de no memoria no siempre está justificada, ya que en tal ejemplo, la ocurrencia del suceso conduce a otro. Y, la tercera hipótesis, también no se logra sostener en varios casos, por ejemplo, cuando ocurren llegadas en grupos, en algunos almacenes, contenedores, entre otros. A pesar de que las tres propiedades no se presentan en distintos problemas de la vida real, son una excelente opción para empezar con el estudio de esos problemas.

Las tres propiedades: homogeneidad, no memoria, y no llegadas simultáneas, combinan en la Definición 1.1 dada en Cinlar (1975).

Definición 1.1. Sea (Ω, B, P) un espacio de probabilidad y sea un proceso $\{N(t), t \geq 0\}$ definido sobre él, de modo que $N(t)$ representa el número de ocurrencias de un suceso en el intervalo $[0, t]$. Diremos que tal proceso es un proceso de Poisson si:

- (i) Para todo $t, s \geq 0$ la distribución de $N(t + s) - N(t)$ es independiente de t . (homogeneidad)
- (ii) Para todo $t, s \geq 0$, $N(t + s) - N(t)$ es independiente de $\{N(u); u \leq t\}$. (no memoria)
- (iii) Las trayectorias muestrales tienen saltos de magnitud unidad. (regularidad)

Nota 1.2. Cuando se usa la definición 1.1, es importante tomar en cuenta las siguientes propiedades:

$$P \{N(t) < \infty\} = 1, \forall t$$

$$E [N(t)] < \infty, \forall t$$

El siguiente resultado cuya demostración incluimos, es clave en el análisis de estos procesos.

Proposición 1.3. Para todo $t \geq 0$ existe una constante $\lambda > 0$ tal que

$$P \{N(t) = 0\} = e^{-\lambda t}$$

Demostración. El número de llegadas en el intervalo $[0, t + s]$ es cero, sii no hay llegadas en el intervalo $[0, t]$ y $(t, t + s]$, de modo que, utilizando la Definición 1.1, la propiedad de *homogeneidad*, la

$$P \{N(t + s) = 0\} = P \{N(t) = 0\} \cdot P \{N(t + s) - N(t) = 0\}$$

Por otro lado, apoyándonos en la propiedad de *no memoria*,

$$P \{N(t + s) - N(t) = 0\} = P \{N(s) = 0\}$$

Por tanto, se define la función $f(t) = P \{N(t) = 0\}$, donde $0 \leq f(t) \leq 1$ y, además, $f(t + s) = f(t) \cdot f(s), \forall s, t \geq 0$. Entonces, $f(t)$ es una función exponencial, o es una función constante igual a cero.

Caso 1:

Supongamos que $f(t) = 0$, para todo $t > 0$, entonces para todo t y casi todo ω , $N(t + s, \omega) - N(t, \omega) \geq 1, \forall s$, esto implica que, el número de llegadas en cualquier intervalo, por pequeño que sea es de al menos una. El número de llegadas en un intervalo $[0, t]$ se puede descomponer como sigue:

$$N(t, \omega) = N(t_1, \omega) + (N(t_2, \omega) - N(t_1, \omega)) + \cdots + (N(t_n, \omega) - N(t_{n-1}, \omega))$$

con $t_1 < t_2 < \dots < t_n = t$, de modo que, se tendría $N(t, \omega) \geq n, \forall n$. Por lo tanto, ocurre que, para casi todo ω , $N(t, \omega) = +\infty, \forall t$, es decir que, las trayectorias muestrales son funciones infinitas, lo cual no tiene sentido. Por lo tanto, se considera el caso 2.

Caso 2:

Se tiene que considerar que, $f(t) = e^{-\lambda t}$, para algún valor de $\lambda \geq 0$. El caso $\lambda = 0$ corresponde al caso degenerado en el que $N(t) = 0, \forall t$.

Proposición 1.4. Si $\{N(t), t \geq 0\}$ es un proceso de Poisson, entonces:

$$(i) \lim_{t \rightarrow 0} \frac{P\{N(t) \geq 2\}}{t} = 0$$

$$(ii) \lim_{t \rightarrow 0} \frac{P\{N(t) = 1\}}{t} = \lambda$$

Demostración. (i) Sea $h(t) = P \{N(t) \geq 2\}$. $N(t, \omega)$, es no decreciente en t , de tal forma que $h(t) \leq h(t + s), \forall t, s$. Sea $n_t = \sup \left\{ n \in \mathbb{N} : n \leq \frac{1}{t} \right\}$, entonces, $t \leq \frac{1}{n_t}$ y $\frac{1}{t} < n_t + 1$, por tanto $h(t) \leq h\left(\frac{1}{n_t}\right)$. Tomando todas estas desigualdades, se tiene que:

$$0 \leq \frac{h(t)}{t} \leq (n_t + 1)h\left(\frac{1}{n_t}\right) = \frac{n_t + 1}{n_t} \cdot n_t h\left(\frac{1}{n_t}\right)$$

Cuando $t \rightarrow 0$, $n_t \rightarrow \infty$, $\frac{(n_t+1)}{n_t} \rightarrow 1$, entonces se debe probar que $n h\left(\frac{1}{n}\right) \rightarrow 0$, mientras $n \rightarrow \infty$, para demostrar que $\frac{h(t)}{t} \rightarrow 0$. Se divide el intervalo $[0, 1]$ en n subintervalos, cada uno con amplitud $\frac{1}{n}$, y, sea S_n la variable número de estos intervalos en los que hay al menos dos llegadas del proceso de Poisson. Esta variable tiene distribución binomial de parámetros n y $p = h(1/n)$, por tanto:

$$E[S_n] = n \cdot h\left(\frac{1}{n}\right)$$

Por la Definición 1.1 (no llegadas simultáneas), cuando n es suficientemente grande, ningún intervalo contiene más de una llegada, de modo que, $\lim_{n \rightarrow \infty} S_n = 0$. Sin embargo, S_n es siempre menor que N_1 , número de llegadas en $[0, 1]$ y $E[N_1] < \infty$ de modo que

$$\lim_{n \rightarrow \infty} E[S_n] = 0.$$

(ii) $P\{N(t) = 1\} = 1 - P\{N(t) = 0\} - P\{N(t) \geq 2\}$, tomando de referencia la proposición 1.3 y aplicando límites se tiene que,

$$\lim_{t \rightarrow \infty} \frac{P\{N(t) = 1\}}{t} = \lim_{t \rightarrow \infty} \frac{1 - e^{-\lambda t}}{t} - \lim_{t \rightarrow \infty} \frac{P\{N(t) \geq 2\}}{t} = \lambda.$$

Para un h suficientemente pequeño, siendo $o(h)$ una función infinitésimo, es decir que $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$, el lema 1.1.3 se escribe como sigue:

- (i) $P\{N(h) \geq 2\} = o(h)$
- (ii) $P\{N(h) = 1\} = \lambda h + o(h)$

A partir de esto, se puede obtener la distribución de probabilidad de las variables del proceso. Por lo que se reformulan los postulados del proceso dando una definición del mismo equivalente a la Definición 1.1.

Definición 1.5. Sea (Ω, B, P) un espacio de probabilidad y sea un proceso $\{N(t), t \geq 0\}$ definido sobre él, de modo que $N(t)$ representa el número de ocurrencias de un suceso en el intervalo $[0, t]$. Diremos que tal proceso es un proceso de Poisson si para h suficientemente pequeño:

- (i) $P\{N(h) = 1\} = \lambda h + o(h), \lambda > 0.$
- (ii) $P\{N(t+h) - N(t) | N(t)\} = P\{N(t+h) - N(t)\}$ para todo $t.$
- (iii) $P\{N(h) \geq 2\} = o(h)$

Esta definición es otra forma de expresar las hipótesis de homogeneidad, no memoria y la de no llegadas simultáneas.

Teorema 1.6. *Si $\{N(t); t \geq 0\}$ es un proceso de Poisson como se define antes, entonces, para todo t , $N(t)$ es una variable aleatoria con distribución de Poisson de parámetro λt , esto es:*

$$P\{N(t) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n = 0, 1, 2, \dots$$

No incluimos la demostración de este resultado.

Para expresar las probabilidades asociadas al proceso usamos la siguiente notación: $P_n(t) = P\{N(t) = n\}, \forall n = 0, 1, \dots, \forall t \geq 0.$

1.1.1. Propiedades básicas del proceso de Poisson homogéneo

Dado que la variable aleatoria $N(t)$ se distribuye según una ley de Poisson de parámetro λt , se tiene que $E[N(t)] = \lambda t$ de ahí que λ se denomine la razón o intensidad del proceso, ya que puede interpretarse como el número medio de llegadas por unidad de tiempo. Además, se puede ver que:

- (i) $E[N(t+s) - N(s)] = \lambda t$, para todo $t \geq 0.$
- (ii) $Var[N(t+s) - N(s)] = \lambda t$, para todo $t \geq 0.$
- (iii) $Cov[N(t+s) - N(s)] = \lambda s$, para todo $t, s \geq 0.$
- (iv) Función generatriz de momentos: $M(t, u) = \exp\{\lambda t(e^u - 1)\}$ para todo $t \geq 0$ y $u \in \mathbb{R}.$
- (v) $P\{N(t+s) - N(t) = k | N(u); u \leq t\} = P\{N(t+s) - N(t) = k\}$

$$= \frac{e^{-\lambda s} (\lambda s)^k}{k!}, k = 0, 1, \dots,$$

Como ya hemos dicho, los postulados de homogeneidad y no memoria frecuentemente son difíciles de verificar en la práctica. Hay situaciones en las que existe con anterioridad una fuerte evidencia que sugiere la independencia de la variable $N(t + s) - N(t)$ con respecto a la historia pasada en cuanto al número de llegadas antes de t , mas, es difícil comprobar la hipótesis sobre la distribución del número de llegadas en un intervalo de tiempo dado. En estos casos es provechosa la siguiente caracterización del proceso de Poisson, cuya mayor ventaja radica en la simplicidad que supone.

Teorema 1.7. $\{N(t); t \geq 0\}$ es un proceso de Poisson si y sólo si:

- Las funciones muestrales del proceso son funciones a saltos de magnitud uno.
- Para todo $t, s \geq 0$, $E[N(t + s) - N(t) | N(u); u \leq t] = \lambda s$

Hay situaciones en las que el axioma de independencia es difícil de comprobar, pero, hay suficientes datos para obtener la distribución del número de llegadas en diversos intervalos de tiempo. En este caso se recomienda tener en cuenta el siguiente teorema, cuya importancia está, en que evita tratar con probabilidades condicionadas. Los resultados que damos a continuación se presentan sin demostración. Para encontrar más detalle puede consultarse el texto de Billingsley (1995).

Teorema 1.8. Un proceso $\{N(t); t \geq 0\}$, es de Poisson de razón λ , si y sólo si, para todo $B \subset \mathbb{R}^+$, obtenido como la unión disjunta de un número finito de intervalos cuya longitud suma b , la distribución de la variable número de llegadas del proceso en el conjunto B es Poisson de parámetro λb .

1.1.2. Distribución de los tiempos entre llegadas y de los tiempos de llegadas

Si se considera un proceso de Poisson $\{N(t); t \geq 0\}$, T_1 es el tiempo de la primera llegada y, T_n es el tiempo transcurrido entre las llegadas $n - 1$ y n , entonces la sucesión $\{T_n; n \geq 1\}$ denota los tiempos entre llegadas del proceso. Interesa determinar la distribución de T_n , para ésto se establece el teorema siguiente:

Teorema 1.9. T_n con $n = 1, 2, \dots$ son independientes e idénticamente distribuidas con distribución exponencial de media $\frac{1}{\lambda}$.

1.1.3. Sucesión de tiempos de espera

La hipótesis de incrementos independientes estacionarios es equivalente a suponer que en cualquier instante de tiempo el proceso se reinicia probabilísticamente: a partir de cualquier punto, el proceso es independiente de todo lo que ha ocurrido previamente y tiene la misma distribución que el proceso original. Es decir, el proceso tiene la propiedad de no memoria, y por tanto es esperable que los tiempos entre llegadas tengan distribución exponencial.

Dado un proceso de Poisson $\{N(t); t \geq 0\}$, se denota S_n al tiempo en que ocurre la n -ésima llegada del proceso, para todo $n = 1, 2, \dots$. $\{S_n, n \geq 1\}$ es una secuencia llamada sucesión de tiempos de espera del proceso, y se puede expresar como

$$S_n = \sum_{i=1}^n T_i, n \geq 1$$

S_n se distribuye según una Erlang de índice n y parámetro λ . La función de densidad que corresponde a una distribución Erlang de índice n y parámetro λ es

$$f_n(t) = e^{-\lambda t} \lambda^n \frac{t^{n-1}}{(n-1)!}$$

Teorema 1.10. *Supongamos un sistema en el que los tiempos de llegada $\{T_n, n = 1, 2, 3, \dots\}$ tienen distribuciones exponenciales idénticas de parámetro λ e independientes. Entonces, el proceso de llegadas asociado es un proceso de Poisson de parámetro λ .*

1.1.4. Distribución condicional de los tiempos de llegadas

Supóngase que durante un tiempo t ha ocurrido una llegada según el proceso de Poisson de parámetro λ , y se desea calcular la distribución del instante de la llegada. Debido a que es un proceso con incrementos estacionarios e independientes, cada subintervalo en $[0, t]$ de la misma amplitud tiene la misma probabilidad de contener tal llegada.

Sea $s \leq t$,

$$P\{T_1 < s | N(t) = 1\} = \frac{P\{T_1 < s, N(t) = 1\}}{P\{N(t) = 1\}}$$

$$= \frac{P_1(s) \cdot P_0(t-s)}{P_1(t)} = \frac{\lambda s e^{-\lambda s} \cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t}$$

La variable T_1 condicionada al suceso $N(t) = 1$, sigue una distribución uniforme en $[0, t]$.

Teorema 1.11. *Dado $N(t) = n$, los tiempos de llegadas $S_1 < S_2 < \dots < S_n$ tienen la misma distribución conjunta que los estadísticos ordenados correspondientes a n variables aleatorias independientes uniformemente distribuidas en $[0, t]$.*

1.1.5. Tiempos residuales

Sea $\{N(t); t \geq 0\}$ un proceso de Poisson con tiempos de llegadas S_1, S_2, \dots . El tiempo de la última llegada antes de t es $S_{N(t)}$ y el tiempo de la siguiente llegada es $S_{N(t)+1}$. Por tanto, dado un tiempo t , el tiempo que ha de transcurrir hasta la siguiente llegada es lo que se va a llamar "tiempo de vida residual" y viene dado por

$$R_t = S_{N(t)+1} - t$$

La distribución de esta nueva variable aleatoria es más sencilla de obtener por las propiedades del proceso de Poisson

$$\begin{aligned} P\{R_t \leq s | N(u), u \leq t\} &= P\{S_{N(t)+1} - t \leq s | N(u), u \leq t\} \\ &= 1 - P\{S_{N(t)+1} > t + s | N(u), u \leq t\} \\ &= 1 - P\{N(t+s) - N(t) = 0 | N(u), u \leq t\} \\ &= 1 - P\{N(s) = 0\} = 1 - e^{-\lambda s} \end{aligned}$$

Es decir que, el tiempo desde t hasta la siguiente llegada es independiente de la historia pasada del proceso hasta t , y además, tiene la misma distribución que cualquier tiempo entre llegadas. No importa el tiempo que haya transcurrido desde la última llegada, la probabilidad de que haya que esperar s unidades de tiempo hasta la siguiente sigue siendo igual. Calculando los valores medios de las variables, se tiene que

$$E[R_t] = \frac{1}{\lambda} = E[S_{n+1} - S_n]$$

1.2. El proceso de Poisson no homogéneo

En el proceso de Poisson se habla de llegadas aleatorias, que no son a tasa constante en el tiempo, es decir que, la función de intensidad es variable en el tiempo t , se está hablando de un Proceso de Poisson no Homogéneo en sus siglas PPNH. Se entiende que un proceso de conteo se basa en definir cuántas veces se ha realizado un evento a lo largo del tiempo; cuando un proceso de conteo se realiza con incrementos independientes y no estacionarios se puede ver como un PPNH.

La siguiente definición dicta los axiomas que indican que un proceso de conteo es un PPNH, Haussmann (2018).

Definición 1.12. Un proceso de conteo $(N_t, t > 0)$ es un Proceso de Poisson No Homogéneo si se cumplen los siguientes axiomas:

- (i) $0 < T_1 < T_2 < \dots < T_n < \dots$; $\lim_{n \rightarrow \infty} T_n = \infty$ y $P\{N(t) < \infty\} = 1$.
- (ii) Para cualquier $t \geq 0$ se cumple que $\lim_{h \rightarrow 0^+} \frac{P(N_{t+h} - N_t \geq 2)}{P(N_{t+h} - N_t = 1)} = 0$
- (iii) Existe una función positiva localmente integrable $\lambda(\cdot)$ definida sobre \mathbb{R}^+ de modo que $\lim_{h \rightarrow 0^+} \frac{1 - P(N_{t+h} - N_t = 0)}{h} = \lambda(t), t \geq 0$

Si se cumplen los axiomas anteriores, resulta que $N_t \sim Poisson(\Lambda(t))$ donde $\Lambda = (\Lambda(t), t \geq 0)$ y $\Lambda(t) = \int_0^t \lambda(s) ds$ para todo $t \geq 0$ es la función de intensidad acumulada.

Definición 1.2.2. Un sistema $\{N(t), t \geq 0\}$ es un PPNH con función de intensidad $\lambda(t)$, si para todo $h \geq 0$,

- (i) $N(0) = 0$
- (ii) Tiene incrementos independientes.
- (iii) $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$
- (iv) $P\{N(t+h) - N(t) \geq 2\} = o(h)$

Se llamará $\Lambda(t) = \int_0^t \lambda(s) ds$, donde $\Lambda(t)$ es una función continua y derivable, cuya derivada es:

$$\frac{d}{dt} \Lambda(t) = \lambda(t)$$

y, $\lambda(t)$ es la función intensidad del proceso.

Teorema 1.13. Sea $\{N(t), t \geq 0\}$ un proceso de Poisson como el que se define en 1.5, entonces:

$$P_n(s, t) = \frac{(\Lambda(s+t) - \Lambda(s))^n}{n!} e^{-(\Lambda(s+t) - \Lambda(s))}$$

Es decir que, $N(s+t) - N(s)$ tiene una distribución Poisson con valor medio $\Lambda(s+t) - \Lambda(s)$.

El PPNH es importante ya que que no se requieren incrementos estacionarios, y, por tanto, se acoge la posibilidad de que las llegadas se produzcan más probablemente en ciertos intervalos de tiempo que en otros.

A partir del teorema 1.13, se deducen las **propiedades 1.2.4**

- (i) $E[N(s+t) - N(s)] = \Lambda(s+t) - \Lambda(s)$
- (ii) $Var[N(s+t) - N(s)] = \Lambda(s+t) - \Lambda(s)$
- (iii) Función generatriz de probabilidades

$$F(t, z) = \exp\{(z-1)[\Lambda(s+t) - \Lambda(s)]\}$$

Una propiedad inmediata que se deduce de la definición del proceso es que $N(t) < n$ si y sólo si $T_n > t$, para todo $n \geq 1$. Una prueba de esta propiedad puede encontrarse en el libro de Rigdon y Basu (2000). A partir de aquí y de la definición del proceso se puede obtener la función de densidad del tiempo de la n -ésima llegada.

Teorema 1.14. El tiempo de la n -ésima ocurrencia de un PPNH con función intensidad $\lambda(t)$ tiene la siguiente función de densidad:

$$f_n(t) = \frac{\lambda(t)\Lambda(t)^{n-1}}{(n-1)!} \exp(-\Lambda(t)), \quad t > 1.$$

Demostración.

$$\begin{aligned} f_n(t) &= -\frac{d}{dt} Pr(T_n > t) = -\frac{d}{dt} Pr(N(t) < n) \\ &= -\frac{d}{dt} \sum_{i=0}^{n-1} Pr(N(t) = i) = -\frac{d}{dt} \sum_{i=0}^{n-1} \frac{\Lambda(t)^i}{i!} \exp(-\Lambda(t)) \\ &= -\sum_{i=0}^{n-1} \frac{d}{dt} \frac{\Lambda(t)^i \exp(-\Lambda(t))}{i!} \end{aligned}$$

$$= \lambda(t) \exp[-\Lambda(t)] \sum_{i=1}^{n-1} \frac{1}{i!} \left\{ \Lambda(t)^i - i\Lambda(t)^{i-1} \right\}$$

Podemos escribir explícitamente los sumandos de la expresión anterior con lo que tenemos:

$$\begin{aligned} f_n(t) &= \lambda(t) \exp(-\Lambda(t)) \left\{ (1 - 0) + (\Lambda(t) - 1) + \left(\frac{1}{2}\Lambda(t)^2 - \Lambda(t)\right) + \dots \right. \\ &\quad \left. + \left(\frac{1}{(n-1)!}\Lambda(t)^{n-1} - \frac{1}{(n-2)!}\Lambda(t)^{n-2}\right) \right\} \\ &= \frac{\lambda(t) \exp(-\Lambda(t)) \Lambda(t)^{n-1}}{(n-1)!}. \end{aligned}$$

El teorema siguiente establece que todo PPNH puede asociarse a un proceso de Poisson homogéneo.

Teorema 1.15. *Sea $\Lambda(t)$ una función continua y no decreciente, entonces T_1, T_2, \dots son los tiempos entre llegadas de un proceso de Poisson no homogéneo, $\{N(t); t > 0\}$ con $\Lambda(t) = E[N(t)]$ si y sólo si $\Lambda(T_1), \Lambda(T_2), \dots$ son los tiempos entre llegadas de un proceso de Poisson homogéneo de razón 1.*

Demostración. Denotemos $\widetilde{N}(t)$ al proceso de recuento asociado a las llegadas $\Lambda(T_1), \Lambda(T_2), \dots$. Usaremos la relación biunívoca entre el número de llegadas hasta t y el tiempo de la n -ésima llegada, para comprobar que para cada $t > 0$ la variable aleatoria $\widetilde{N}(t)$ tiene distribución de Poisson con media t , lo que demuestra el resultado. Denotamos Λ^{-1} a la función inversa de la función Λ , entonces

$$Pr(\widetilde{N}(t) < n) = Pr(\Lambda(T_n) > t) = Pr(T_n > \Lambda^{-1}(t)) = Pr(N(\Lambda^{-1}(t)) < n)$$

donde la segunda igualdad se obtiene porque Λ es una función estrictamente creciente.

$$\begin{aligned} Pr(N(\Lambda^{-1}(t)) < n) &= \sum_{i=1}^{n-1} \frac{(\Lambda(\Lambda^{-1}(t)))^i}{i!} \exp(-\Lambda(\Lambda^{-1}(t))) \\ &= \sum_{i=1}^{n-1} \frac{t^i}{i!} e^{-t}. \end{aligned}$$

Este resultado es la base para la simulación de trayectorias del proceso de Poisson que implementamos en el capítulo siguiente.

1.3. Algunas propiedades del PPNH

En esta sección incluimos algunos resultados sobre la estructura del PPNH que serán útiles en el capítulo siguiente en la estimación máximo-verosímil. Es importante distinguir dos situaciones en la observación del proceso. En el primer caso consideremos que la observación finaliza cuando se ha registrado un predeterminado número de fallos, los datos se dicen *truncados en número de fallos*. En el segundo caso consideramos que la observación finaliza pasado un determinado periodo de tiempo fijado de antemano. Los datos se denominan en este caso *truncados en tiempo*.

1.3.1. Truncamiento en número de fallos

Obtenemos a continuación la función de verosimilitud conjunta de los tiempos de fallo registrados: $T_1 < T_2 < \dots, < T_n$.

Teorema 1.16. *La función de densidad conjunta de los tiempos de fallo $T_1 < T_2 < \dots, < T_n$ de un PPNH con función intensidad $\lambda(t)$ es*

$$f(t_1, t_2, \dots, t_n) = \left(\prod_{i=1}^n \lambda(t_i) \right) \exp \left(- \int_0^{t_n} \lambda(x) dx \right)$$

Demostración. Usamos para esta demostración la siguiente propiedad:

$$f(t_1, t_2, \dots, t_n) = f(t_1) f(t_2|t_1) f(t_3|t_2, t_1) \cdots f(t_n|t_{n-1}, \dots, t_2, t_1),$$

para toda secuencia $t_1 < t_2 < \dots < t_n$. Usando la equivalencia entre los sucesos $N(t) = 0$ y $T_1 > t$, tenemos, de forma análoga al caso homogéneo la distribución del primer tiempo de llegada:

$$S_1(t_1) = Pr(T_1 > t_1) = Pr(N(0, t_1] = 0) = \exp \left(- \int_0^{t_1} \lambda(s) ds \right).$$

La correspondiente función de densidad se obtiene derivando la expresión anterior como $f_1(t_1) = \lambda(t_1) \exp \left(- \int_0^{t_1} \lambda(s) ds \right)$. Con respecto a la variable condicionada T_2 condicionada al suceso $T_1 = t_1$, podemos escribir

$$\begin{aligned} S_2(t_2|t_1) &= Pr(T_2 > t_2|T_1 = t_1) = Pr(N(t_1, t_2] = 0) \\ &= \exp \left(- \int_{t_1}^{t_2} \lambda(s) ds \right), t_2 > t_1 > 0. \end{aligned}$$

La correspondiente densidad condicional es por tanto

$$f_2(t_2|t_1) = \lambda(t_2) \exp \left(- \int_{t_1}^{t_2} \lambda(s) ds \right), t_2 > t_1 > 0.$$

Usando la propiedad de no memoria del proceso de Poisson, obtenemos que la función de supervivencia condicional de la variable T_k dada la historia del proceso se puede simplificar

$$\begin{aligned} S_k(t_k|t_1, t_2, \dots, t_{k-1}) &= Pr(T_k > t_k | T_{k-1} = t_{k-1}) = Pr(N(t_{k-1}, t_k] = 0) \\ &= \exp\left(-\int_{t_{k-1}}^{t_k} \lambda(s) ds\right), t_k > t_{k-1} > 0. \end{aligned}$$

Así

$$f_k(t_k|t_{k-1}) = \lambda(t_k) \exp\left(-\int_{t_{k-1}}^{t_k} \lambda(s) ds\right), t_k > t_{k-1}.$$

La función de densidad conjunta de T_1, T_2, \dots, T_n , se expresa

$$\begin{aligned} f(t_1, t_2, \dots, t_n) &= \left[\lambda(t_1) \exp\left(-\int_0^{t_1} \lambda(s) ds\right) \right] \left[\lambda(t_2) \exp\left(-\int_{t_1}^{t_2} \lambda(s) ds\right) \right] \\ &\quad \times \dots \times \left[\lambda(t_k) \exp\left(-\int_{t_{k-1}}^{t_k} \lambda(s) ds\right) \right]. \end{aligned}$$

Reordenando los términos se obtiene el resultado deseado.

La distribución condicional de los $n - 1$ tiempos de fallo dado el n -ésimo tiempo de fallo puede obtenerse como se demuestra en el siguiente resultado. Este resultado es la generalización del resultado análogo (ya enunciado anteriormente) que, para el proceso homogéneo, establece que dado que ha ocurrido el suceso $N(t) = n$, la distribución de los tiempos de las n llegadas coincide con la distribución de los n primeros estadísticos ordenados asociados a la distribución Uniforme en el intervalo $(0, t)$.

Teorema 1.17. *Si los tiempos de fallo de un proceso de Poisson no homogéneo son $T_1 < T_2 < \dots, T_n$, entonces, condicionado al suceso $T_n = t_n$, las variables aleatorias $T_1 < T_2 < \dots, T_{n-1}$ se distribuyen como los $n - 1$ estadísticos ordenados de la distribución con función de distribución acumulada dada por*

$$G(y) = \begin{cases} 0, & y \leq 0; \\ \Lambda(y)/\Lambda(t_n), & 0 < y \leq t_n; \\ 1, & y > t_n. \end{cases}$$

Demostración. La función de densidad conjunta de los estadísticos ordenados a partir de la distribución $G(y)$ es, para $0 < y_1 < y_2 < \dots < y_n$,

$$f(y_1, y_2, \dots, y_n) = (n - 1)! \prod_{i=1}^{n-1} G'(y_i) = (n - 1)! \prod_{i=1}^{n-1} \frac{\Lambda(y_i)}{\Lambda(t_n)},$$

Usando los resultados probados en los teoremas 1.2.4 y 1.3.1, obtenemos la siguiente distribución

$$\begin{aligned} f(t_1, t_2, \dots, t_{n-1} | t_n) &= \frac{f(t_1, t_2, \dots, t_{n-1}, t_n)}{f(t_n)} = \frac{(\prod_{i=1}^n \lambda(t_i)) \exp(-\Lambda(t_n))}{\frac{\lambda(t_n)(\Lambda(t_n))^{(n-1)}}{(n-1)!} \exp(-\Lambda(t_n))} \\ &= \frac{(n-1)! (\prod_{i=1}^{n-1} \lambda(t_i))}{(\Lambda(t_n))^{(n-1)}} = (n-1)! \prod_{i=1}^{n-1} \frac{\lambda(t_i)}{\Lambda(t_n)}. \end{aligned}$$

Esta última expresión es la distribución conjunta de los $n - 1$ estadísticos ordenados de una distribución dada por $G(y)$.

1.3.2. Truncamiento en tiempo de fallos

En este caso la observación del proceso finaliza cuando ha transcurrido un tiempo previamente especificado. En este caso, el número de fallos registrados hasta ese tiempo τ , es decir $N(\tau)$ es una variable aleatoria. Esta aleatoriedad debe tenerse en cuenta a la hora de determinar la función de verosimilitud.

Un resultado paralelo al Teorema 1.3.2 es el siguiente.

Teorema 1.18. *Si un PPNH con intensidad $\lambda(t)$ es observado hasta el tiempo τ , y se registran tiempos de fallo $T_1 < T_2 \dots < T_n < \tau$. Suponiendo que el número total de fallos registrados en $(0, \tau]$ es $N(\tau) = n$, condicionado a este suceso, las variables aleatorias $T_1 < T_2 \dots < T_n$ se distribuyen como los estadísticos ordenados de una variable con función de distribución dada por*

$$G(y) = \begin{cases} 0, & y \leq 0; \\ \Lambda(y)/\Lambda(\tau), & 0 < y \leq \tau; \\ 1, & y > \tau. \end{cases}$$

Demostración. La función de densidad conjunta de $T_1 < T_2 \dots < T_n$ dado $N(\tau) = n$ es

$$\begin{aligned} f(t_1, \dots, t_n | n) &= \lim_{\epsilon_i \rightarrow 0} \frac{Pr(t_1 < T_1 < t_1 + \epsilon_1, \dots, t_n < T_n < t_n + \epsilon_n | n)}{\prod_{i=1}^n \epsilon_i} \\ &= \lim_{\epsilon_i \rightarrow 0} \frac{Pr(t_1 < T_1 < t_1 + \epsilon_1, \dots, t_n < T_n < t_n + \epsilon_n, N(\tau) = n)}{Pr(N(\tau) = n) \prod_{i=1}^n \epsilon_i} \end{aligned}$$

Puesto que las cantidades ϵ_i son infinitesimales, puede asumirse que $t_i + \epsilon_i < t_{i+1}$, para todo i , de modo que la expresión en el numerador puede re-escribirse considerando que los fallos se distribuyen en el intervalo $(0, \tau]$, como indica la expresión a continuación

$$Pr(N(0, t_1] = 0) Pr(N(t_1, t_1 + \epsilon_1] = 1) Pr(N(t_1 + \epsilon_1, t_2] = 0)$$

$$\begin{aligned} &Pr(N(t_2, t_2 + \epsilon_2] = 1) \times \dots \times \\ &Pr(N(t_n, t_n + \epsilon_n] = 1)Pr(N(t_n + \epsilon_n, \tau] = 0) \end{aligned}$$

Sustituyendo cada factor según el valor de la correspondiente distribución de Poisson, es decir, de parámetro $\Lambda(t_i + \epsilon_i) - \Lambda(t_i)$, para cada intervalo $(t_i, t_i + \epsilon_i]$ o bien $\Lambda(t_{i+1}) - \Lambda(t_i + \epsilon_i)$, en los intervalos de tipo $(t_i + \epsilon_i, t_{i+1}]$, y agrupando convenientemente los términos obtenidos, tenemos lo siguiente

$$\begin{aligned} f(t_1, \dots, t_n | n) = & \\ e^{-\Lambda(\tau)} & \left(\lim_{\epsilon_1 \rightarrow 0} \frac{\Lambda(t_1 + \epsilon_1) - \Lambda(t_1)}{\epsilon_1} \right) \times \dots \times \\ & \left(\lim_{\epsilon_n \rightarrow 0} \frac{\Lambda(t_n + \epsilon_n) - \Lambda(t_n)}{\epsilon_n} \right) / \left(\frac{\exp(-\Lambda(\tau))(\Lambda(\tau))^n}{n!} \right) \end{aligned}$$

Cada uno de estos límites es la derivada de Λ , de modo que finalmente obtenemos

$$f(t_1, \dots, t_n | n) = n! \prod_{i=1}^n \frac{\lambda(t_i)}{\Lambda(\tau)},$$

siendo esta última expresión la distribución conjunta de los n estadísticos ordenados de una muestra de tamaño n de la distribución $G(y)$.

1.4. El PPNH en Fiabilidad. La función ROCOF

En esta sección analizamos las características más importantes del proceso de Poisson como modelo de ocurrencia de fallos en análisis de fiabilidad.

1.4.1. El concepto de reparación mínima

Los PPNHs son buenos modelos para sistemas sujetos a fallos y reparaciones. Una vez que se ha producido el fallo en el sistema, se llevan a cabo ciertas actividades de mantenimiento y el sistema vuelve a estar operativo. Para formular un modelo de evolución del sistema es esencial describir las características del sistema una vez que se ha completado la acción de mantenimiento. Los sistemas complejos suelen repararse en lugar de sustituirse tras un fallo. En la práctica, es más conveniente, sobre todo desde el punto de vista económico, restablecer la funcionalidad del sistema por métodos distintos a la sustitución de todo el sistema. En particular, muchas reparaciones en la

vida real devuelven el sistema a condiciones operativas que son básicamente las mismas que había antes de que se produjera el fallo. Como se señala en Aven (2007), el objetivo es devolver el sistema a un estado de funcionamiento lo antes posible y, por tanto, parece razonable que, tras la reparación, el sistema vuelva a estar exactamente en el mismo estado que tenía antes del fallo. Nos referimos a esta situación diciendo que el sistema está en "tan mal como antes", lo que nos lleva al concepto de Reparación Mínima introducido por Ascher (1968) donde se considera que la edad (en el sentido de eficacia) del sistema tras la reparación es la misma que tenía antes del fallo.

La situación anterior se da con frecuencia en el mantenimiento de sistemas complejos (con muchos componentes), en los que, tras un fallo, se sustituyen pequeñas partes del sistema por otras nuevas. En este caso, las reparaciones o sustituciones perfectas se limitan al nivel de los componentes. Si adoptamos la formulación de Ascher (véase Ascher, 2007a), se entiende que un sistema es un elemento reparable, mientras que un componente del sistema se considera no es reparable y, por tanto, cuando un componente del sistema falla, se sustituye por uno nuevo. De este modo, si un componente de un sistema complejo falla y se sustituye, se puede considerar que el sistema en su conjunto está aproximadamente en el mismo estado que inmediatamente antes del fallo.

Al tratarse de sistemas reparables, consideramos que el sistema sufre múltiples fallos a lo largo del tiempo, por lo que necesitamos un modelo estocástico que describa la ocurrencia de eventos (que son fallos del sistema) en el tiempo, es decir, un proceso puntual. En general, el tiempo transcurrido entre los sucesivos fallos no es independiente ni está idénticamente distribuido, como ocurre con los modelos de Reparación Perfecta, de los que un caso particular es el proceso de Poisson homogéneo considerado anteriormente.

Considere un sistema reparable que se pone en funcionamiento por primera vez en $S_0 = 0$, sea $S_1 < S_2 < S_3 < \dots$, los tiempos aleatorios ordenados en los que se producen los sucesivos fallos. Ignoramos el tiempo de duración de las reparaciones, por lo que suponemos que una vez que el sistema falla, se pone inmediatamente en funcionamiento. Los tiempos de interlegada se denotarán como $T_i = S_i - S_{i-1}$, para $i \geq 1$. Se necesitará un modelo estocástico para la evolución del sistema y deberá especificar adecuadamente la distribución conjunta de $\{T_1, T_2, \dots\}$.

El proceso estocástico de recuento correspondiente es $\{N(t); t \geq 0\}$ donde $N(t)$ denota el número de fallos que se producen en el intervalo $(0, t]$. Así $N(t_1, t_2) = N(t_2) - N(t_1)$ cuenta el número de fallos que se han producido en el intervalo $(t_1, t_2]$. Se observa que (con probabilidad 1) $N(t)$ es una función escalonada no decreciente y continua hacia la derecha. Una referencia básica en el tratamiento de los procesos de conteo es Andersen et al. (1993). Véase también Rausand y Høyland (2004) para una introducción a la teoría.

Las siguientes definiciones introducen algunas funciones clave para describir la evolución de los sistemas reparables.

Definición 1.19. Función media de un proceso de recuento

Sea $\{N(t); t \geq 0\}$ un proceso de conteo, la función media del proceso se define como la esperanza

$$\Lambda(t) = E[N(t)], \quad (1.1)$$

para $t \geq 0$.

Por tanto, $\Lambda(t)$ es el número esperado de fallos hasta el tiempo t , y es una función no decreciente de t . También puede verificarse que $\Lambda(t)$ es una función continua hacia la derecha (véase, por ejemplo, Rigdon y Basu, 2000). La siguiente definición da una medida importante que caracteriza algunos tipos de procesos de recuento.

Definición 1.20. Tasa de ocurrencia de fallos (ROCOF) Cuando la función media $\Lambda(t)$ es diferenciable, se puede definir la siguiente función

$$m(t) = \frac{d}{dt}\Lambda(t). \quad (1.2)$$

para $t \geq 0$.

Esta función suele interpretarse como la tasa de variación instantánea del número medio de fallos.

A continuación se presenta una función estrechamente relacionada, que ya hemos definido.

Definición 1.21. Función de intensidad incondicional

Sea $\{N(t); t \geq 0\}$ un proceso de conteo, la **función de intensidad (incondicional)** viene dada por

$$\lambda(t) = \lim_{\epsilon \rightarrow 0} \frac{P\{N(t+\epsilon) - N(t) = 1\}}{\epsilon}. \quad (1.3)$$

La probabilidad de fallo en un intervalo pequeño puede aproximarse mediante el producto de la función de intensidad y la longitud del intervalo. En un intervalo donde $\lambda(t)$ es grande, se esperan muchos fallos, mientras que en los intervalos donde $\lambda(t)$ es pequeño, se esperan pocos fallos.

Cuando la probabilidad de fallos simultáneos es cero, es decir, cuando el proceso es ordenado, se puede demostrar (véase Rigdon y Basu, 2000) que la función de intensidad es igual al ROCOF, es decir, $\lambda(t) = m(t)$.

Cabe destacar que, a diferencia de la función de riesgo de una variable aleatoria, por ejemplo $\lambda(t)$, la función de intensidad no se define en términos de una probabilidad condicional. En otras palabras, la función de intensidad es la probabilidad incondicional de un fallo en un intervalo pequeño dividida por la longitud del intervalo. Evidentemente, se trata de una medida asociada a los sistemas reparables, mientras que la función de peligro no lo es. En la definición de la función de intensidad, no se dice nada sobre la historia del sistema hasta t , el único interés es la ocurrencia del siguiente fallo, que no es necesariamente el primero. En la definición de la función de peligro, la historia del sistema (no fallado) hasta el momento actual aparece implícita, ya que se está calculando la probabilidad de fallo (ahora el primero) dado que el sistema aún no ha fallado.

Mientras que la función media $\Lambda(t)$, el ROCOF $m(t)$ y la intensidad incondicional $\lambda(t)$ son conceptos útiles en la descripción de un proceso de conteo, no caracterizan completamente las propiedades estocásticas del proceso. En su lugar, es necesario un índice de información más completo, dado en la siguiente definición.

Definición 1.22. Función de intensidad condicional

Dejemos que \mathcal{F}_{t-} denote la historia del proceso hasta, pero sin incluir, el tiempo t . \mathcal{F}_{t-} está generado por el conjunto $\{N(s), 0 \leq s < t\}$, y por tanto contiene toda la información sobre los tiempos de fallo en el pasado, es decir hasta t , y se llama a *filtración*. La función de intensidad condicional se define como

$$\gamma(t) = \lim_{\epsilon \rightarrow 0} \frac{P \{N(t + \epsilon) - N(t) \geq 1 \mid \mathcal{F}_{t-}\}}{\epsilon}. \quad (1.4)$$

Esta función no debe confundirse, en general, con el ROCOF $m(t)$ dado en la definición 1.20, ni, por supuesto, con la función de intensidad incondicional $\lambda(t)$. Sin embargo, en el caso particular de los PPNH, como veremos a continuación, la intensidad condicional depende de la historia del proceso sólo a través del tiempo t , es decir, a diferencia de otros modelos más complejos (como los procesos de Cox, los procesos de renovación o los modelos de intensidad multiplicativa, entre otros), para un PPNH, $\gamma(t)$ es una función (determinista) sólo del tiempo t y se cumple que $\gamma(t) = m(t)$.

1.4.2. El modelo: Proceso de Poisson no homogéneo (PPNH)

Consideremos un sistema reparable, con la notación de la sección anterior. Denotemos F como la función de distribución acumulativa correspondiente

a $T = T_1$ (tiempo hasta que el sistema falla por primera vez); $\bar{F} = 1 - F$ la función de supervivencia o función de fiabilidad y $\lambda(t)$ la función de riesgo, entonces

$$\bar{F}(t) = \exp \left[- \int_0^t h(u) du \right]. \quad (1.5)$$

A continuación damos la definición matemática de Reparación Mínima, que fue introducida por Ascher (1968). Véase Hollander et al. (2007) o Finkelstein (2008) para referencias recientes.

Definición 1.23. Reparación mínima La función de supervivencia de un sistema que ha fallado y ha sido *instantáneamente reparado mínimamente* a la edad s se define como

$$\bar{F}_s(t) = \frac{1 - F(s+t)}{1 - F(s)} = \exp \left[- \int_s^{s+t} h(u) du \right]. \quad (1.6)$$

De la definición anterior se deduce que, dado que el tiempo observado del primer fallo es $S_1 = s_1$ la función de supervivencia correspondiente al tiempo de espera al segundo fallo bajo Reparación Mínima viene dada por \bar{F}_{s_1} . En general, cuando se realizan reparaciones mínimas en el sistema, dado que el fallo j th se ha producido en el momento $S_j = s_j$, tenemos que

$$P \{ T_{j+1} > t | S_j = s_j \} = \bar{F}_{s_j}(t), \quad \text{for } t \geq 0, \quad j = 1, 2, \dots \quad (1.7)$$

es la función de supervivencia correspondiente al siguiente tiempo de llegada.

Veremos que la propiedad definida anteriormente de Reparación Mínima es satisfecha por el Proceso de Poisson No Homogéneo (PPNH) definido en la Sección 1.2. Este proceso se ha aplicado ampliamente en la ingeniería de la fiabilidad, en particular en el estudio de la modelización de la fiabilidad del software (véase Pham y Zhang (2003) y sus referencias).

Sea $\lambda(t)$ la función intensidad del proceso de Poisson dado por $\{N(t); t \geq 0\}$, y sea $\Lambda(t)$ la función media asociada, tal como se define en la expresión (1.1). Entonces se puede deducir que para cualquier $t \geq 0$ la variable aleatoria $N(t)$ sigue una distribución de Poisson con parámetro $\Lambda(t)$, es decir, la probabilidad de que $N(t)$ sea un número entero dado n se expresa como

$$P\{N(t) = n\} = \frac{(\Lambda(t))^n}{n!} e^{-\Lambda(t)}; \quad n = 0, 1, 2, \dots \quad (1.8)$$

Con $\Lambda(t)$ y $\lambda(t)$ definidos en el apartado anterior. En general, la probabilidad de que el número de fallos en un intervalo $(s, s+t]$, para cualquier $s, t > 0$ sea un determinado n se expresa como

$$P\{N(s, s+t) = n\} = \frac{[\Lambda(s+t) - \Lambda(s)]^n}{n!} \exp [-(\Lambda(s+t) - \Lambda(s))], \quad (1.9)$$

para $n = 0, 1, 2, \dots$

Se puede caracterizar $\{N(t); t \geq 0\}$ como una lista ordenada $S_1 < S_2 < \dots$ de tiempos de eventos. El proceso estocástico se especifica completamente por la distribución conjunta de $N(t_1) \leq N(t_2) \leq \dots \leq N(t_n)$ para cualquier secuencia $t_1 \leq t_2 \leq \dots \leq t_n$ y $n \geq 0$. Esto equivale a especificar la distribución conjunta de la secuencia $\{S_1 < S_2 < \dots\}$. A partir de las expresiones (1,8) y (1,9) y teniendo en cuenta la propiedad de los incrementos independientes, podemos obtener lo siguiente

$$\begin{aligned} P\{S_1 > s\} &= P\{N(s) = 0\} = e^{-\Lambda(s)} = \\ &= \exp\left[-\int_0^s m(u)du\right] = \exp\left[-\int_0^s h(u)du\right]. \end{aligned} \quad (1.10)$$

Así, el primer evento del proceso es una variable aleatoria con $T_1 = S_1$ con función de riesgo $h(t) = \lambda(t)$. Denotemos con F su correspondiente función de distribución. Además,

$$\begin{aligned} P\{S_2 > s_1 + t | S_1 = s_1\} &= P\{N(s_1, s_1 + t) = 0\} = \\ &= e^{-(\Lambda(s_1+t) - \Lambda(s_1))} = \exp\left[-\int_{s_1}^{s_1+t} \lambda(u)du\right], \end{aligned} \quad (1.11)$$

y, en general nosotros tenemos

$$\begin{aligned} P\{S_{j+1} > s_j + t | S_j = s_j\} &= P\{N(s_j, s_j + t) = 0\} = \\ &= e^{-(\Lambda(s_j+t) - \Lambda(s_j))} = \exp\left[-\int_{s_j}^{s_j+t} \lambda(u)du\right]. \end{aligned} \quad (1.12)$$

Así, según la notación de la Definición 1.23 y la ecuación (1,6), se puede deducir que la distribución de los tiempos de interlegada de un PPNH que son $T_j = S_j - S_{j-1}$ for $j = 1, 2, \dots$ y $S_0 = 0$, se expresa de la siguiente manera

$$P\{T_{j+1} > t | S_j = s_j\} = \bar{F}_{s_j}(t), \quad (1.13)$$

donde F , como en el caso anterior, es la función de distribución del tiempo del primer evento del PPNH, es decir $T_1 = S_1$.

La conclusión inmediata es que el PPNH satisface la propiedad que define la Reparación Mínima en (1,7). Por lo tanto, los tiempos de los eventos de un PPNH pueden interpretarse como tiempos en los que el sistema se repara mínimamente. Además, la función de intensidad condicional $\gamma(t)$ del proceso de recuento de fallos es igual a la función de intensidad incondicional $\lambda(t)$,

que también es el ROCOF $m(t)$, y también da la función de peligro de la variable aleatoria que representa el primer tiempo hasta el fallo del sistema T_1 , por lo que tenemos

$$\gamma(t) = \lambda(t) = m(t) = h(t). \quad (1.14)$$

En lo sucesivo, utilizaremos para esta función la notación $\lambda(t)$.

Se podría pensar que los modelos no homogéneos son mucho menos manejables que los modelos estacionarios. Sin embargo, la dificultad inherente a la función de renovación $\Lambda(t)$ es notable, es decir, el número esperado de fallos en un intervalo finito $(0, t]$, que debe calcularse cuando se realizan reparaciones perfectas en el sistema. Como sabemos, en este caso, $\Lambda(t)$ viene dada por una suma de infinitas funciones de convolución, $\Lambda(t) = \sum_{k=0}^{\infty} F^{(k)}(t)$, donde F puede entenderse como la función de distribución del tiempo hasta el primer fallo. Entre otras consideraciones, en el capítulo se explicó que la función de renovación no está disponible en forma cerrada, salvo en casos especiales como, por ejemplo, cuando F representa una ley exponencial. La función análoga cuando el sistema se repara mínimamente, es decir, cuando el modelo es un PPNH, es lo que llamamos llamada función media, y puede obtenerse como $\Lambda(t) = \int_0^t \lambda(u) du$ la función de riesgo acumulativo del tiempo aleatorio hasta el primer fallo, véase Ascher (2007b) para más detalles.

La distribución de un PPNH y, en consecuencia, el comportamiento de la fiabilidad a lo largo del tiempo de un sistema mínimamente reparado, esta se caracteriza completamente por la función ROCOF. La forma de la ROCOF puede mostrar muchos tipos de comportamiento, por ejemplo, decreciente (los sistemas con esta propiedad se denominan a veces “felices”), creciente (para sistemas de deterioro), en forma de bañera, en forma de bañera invertida por mencionar algunas, pero las posibilidades son innumerables y podrían incluir un comportamiento cíclico con múltiples periodicidades o marcadas tendencias asimétricas. En este sentido, es interesante el trabajo de Krivtsov (2007), que considera alternativas a los dos modelos tradicionalmente y más ampliamente utilizados y que se denominan modelo *log-lineal* y modelo *power law*. Propone el uso de las formas paramétricas correspondientes a un gran número de familias de distribución tradicionales para la función de distribución subyacente del proceso. Sin embargo, los modelos paramétricos no suelen ser lo suficientemente flexibles como para cubrir todas las formas posibles de la función de intensidad y, por tanto, los procedimientos no paramétricos resultan muy atractivos. En concreto, el objetivo principal de este trabajo es estimar la función λ mediante técnicas no paramétricas a partir de una o varias realizaciones de un PPNH.

Capítulo 2

Inferencia paramétrica en el PPNH

2.1. Datos de recurrencia en fiabilidad

Los procesos puntuales se utilizan como modelos para representar datos de recurrencia, estos datos de recurrencia emergen continuamente en aplicaciones de fiabilidad. Entre las aplicaciones más importantes se encuentran las de *reparación de sistemas*. El proceso de reparación de un sistema individual se representa como una secuencia de tiempos de reparación T_1, T_2, T_3, \dots . El término *reparación de sistemas* describe ya sea un ajuste, un fallo, una sustitución o un evento de interés general. Las reparaciones se observan comúnmente a lo largo de un intervalo de observación fijo (t_0, t_a) donde, normalmente, $t_0 = 0$. Las aplicaciones pueden tener datos sobre un solo sistema, de una muestra o de una colección de sistemas. Si los datos de una colección se combinan en datos para formar un único proceso, el proceso que resulta se conoce como *proceso puntual superpuesto* o *superposición de varios procesos puntuales*.

Los datos de los sistemas reparables se recolectan para estimar o predecir cantidades como: la distribución de tiempos entre reparaciones; el número acumulado de reparaciones en el intervalo $(0, t]$ en función de la edad t del sistema; el tiempo esperado entre fallos, conocido también como tiempo medio entre fallos *MTBF*; el número esperado de reparaciones en el intervalo $(0, t]$ en función de t ; la tasa de reparación en función de t ; y, el costo medio de las reparaciones en función de t .

Para un sistema único, los datos de recurrencia pueden expresarse como $N(s, t)$, el número acumulado de recurrencias en el intervalo de *edad del sistema* $(s, t]$. Cuando se representa a $N(0, t)$, se utiliza la notación $N(t)$. El

modelo que describe una población de sistemas, está basado en la función acumulativa media *MCF* a la edad t del sistema; esta función de la población, *MCF*, es definida como $\Lambda(t) = E[N(t)]$, donde la esperanza es sobre la variabilidad de cada sistema y la variabilidad unidad en la población. Supongamos que $\Lambda(t)$ es diferenciable, entonces $\lambda(t)$ define la tasa de recurrencia por sistema para la población,

$$\lambda(t) = \frac{dE[N(t)]}{dt} = \frac{d\Lambda(t)}{dt}$$

Como se menciona en la sección 2.1, para los datos de recurrencia son útiles los modelos paramétricos de procesos puntuales. Entre los modelos más utilizados para datos de recurrencia están: los procesos de Poisson homogéneos y los no homogéneos, los procesos de renovación y versiones superpuestas de estos procesos.

El proceso de Poisson es un modelo paramétrico simple que se utiliza habitualmente para analizar cierto tipo de datos de recurrencia.

Como ya hemos visto antes en este trabajo, el proceso de Poisson homogéneo (PPH) es un proceso de Poisson con una tasa de recurrencia constante, por ejemplo $\lambda(t) = \frac{1}{\theta}$. En este caso: $N(a, b)$ tiene una distribución de Poisson con parámetro $\Lambda(a, b) = \frac{b-a}{\theta}$; el número esperado de recurrencias en $(a, b]$ es $\Lambda(a, b)$, equivalentemente, el número esperado de recurrencias por unidad de tiempo sobre $(a, b]$ es constante e igual a $\frac{1}{\theta}$, esta propiedad se llama *incrementos estacionarios*; los tiempos entre recurrencias $\tau_j = T_j - T_{j-1}$ son independientes e idénticamente distribuidos (iid), cada uno con una distribución *EXP*(θ).

Por el contrario, un proceso de Poisson no homogéneo (PPNH) es un modelo con una tasa de recurrencia no constante $\lambda(t)$ (ver Capítulo 1). En este caso, los tiempos de interrecurrencia no son independientes ni idénticamente distribuidos. El número esperado de recurrencias por unidad de tiempo sobre $(a, b]$ es $\frac{\Lambda(a, b)}{b-a} = \frac{1}{b-a} \int_a^b \lambda(t) dt$. Un modelo PPNH suele especificarse en términos de la tasa de recurrencia $\lambda(t)$. Para especificar un modelo PPNH se utiliza $\lambda(t) = \lambda(t; \theta)$, una función de un vector desconocido de parámetros θ . Por ejemplo, la tasa de recurrencia del modelo de potencia es $\lambda(t; \beta, \eta) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$, $\beta > 0, \eta > 0$. El correspondiente número medio acumulado de recurrencia sobre $(0, t]$ es $\Lambda(t; \beta, \eta) = \left(\frac{t}{\eta}\right)^\beta$. Cuando $\beta = 1$ (ver teorema 1.2.4), este modelo se reduce al modelo PPH. La tasa de recurrencia del modelo loglineal es $\lambda(t; \gamma_0, \gamma_1) = \exp(\gamma_0 + \gamma_1 t)$; el correspondiente número acumulado de recurrencias sobre $(0, t]$ es $\Lambda(t; \gamma_0, \gamma_1) = \frac{[\exp(\gamma_0)][\exp(\gamma_1 t) - 1]}{\gamma_1}$. Cuando $\gamma_1 = 0$, $\lambda(t; \gamma_0, 0) = \exp(\gamma_0)$, el modelo es un PPH.

Una secuencia de recurrencias del sistema en los tiempos T_1, T_2, \dots , es un

proceso de renovación si los tiempos de interrecurrencia $\tau_j = T_j - T_{j-1}, j = 1, 2, \dots, (T_0 = 0)$ son idénticamente distribuidos e independientes. La función acumulativa media, MCF, de un proceso de renovación también se conoce como *función de renovación*. Un proceso de Poisson homogéneo puede verse como un proceso de renovación tal que los tiempos de interrecurrencia $\tau_j \sim EXP(\theta)$ con $\theta = \frac{1}{\lambda}$.

El proceso puntual que surge de la agregación de las renovaciones de un grupo de procesos de renovación independientes que operan simultáneamente se conoce como proceso de renovación superpuesto (SRP). A menos que los procesos de renovación individuales sean PPH, un SRP no es un proceso de renovación. Sin embargo, el teorema de Drenick (Drenick, 1960) dice que cuando el número de sistemas n es grande y los sistemas han funcionado lo suficiente como para eliminar transitorios, un SRP se comporta como un PPH. Este resultado es más bien un teorema del límite central para los procesos de renovación. Este resultado se utiliza a veces para justificar el uso de la distribución exponencial para modelar los tiempos de interrecurrencia en una gran población de sistemas. Dependiendo de la forma de las distribuciones subyacentes en los procesos individuales de renovación superpuestos, pueden ser necesarias muestras grandes y tiempos largos para que la aproximación PPH sea adecuada (véase Blumenthal, Greenwood, Herbach 1973, 1976). En la práctica, se puede comprobar la adecuación de la aproximación mediante la simulación.

2.2. Estimación máximo verosimil en PPNH

En esta sección revisamos los estimadores de máxima-verosimilitud para los modelos paramétricos más usados: el modelo con tasa de fallo potencial y el modelo con tasa de fallo log-lineal.

2.2.1. Función de verosimilitud en el proceso de Poisson

Para un sistema que ha sido observado durante un período $(0, t_a]$, los datos son el número de recurrencias d_1, d_2, \dots, d_m en los intervalos no solapados $(t_0, t_1], (t_1, t_2], \dots, (t_{m-1}, t_m]$; $(t_0 = 0, t_m = t_a)$. La probabilidad para el modelo PPNH es

$$L(\theta) = Pr [N(t_0, t_1) = d_1, \dots, N(t_{m-1}, t_m) = d_m] = \prod_{j=1}^m Pr [N(t_{j-1}, t_j) = d_j]$$

$$\begin{aligned}
&= \prod_{j=1}^m \frac{[\Lambda(t_{j-1}, t_j; \theta)]^{d_j}}{d_j!} \exp[-\Lambda(t_{j-1}, t_j; \theta)] \\
&= \prod_{j=1}^m \frac{[\Lambda(t_{j-1}, t_j; \theta)]^{d_j}}{d_j!} \times \exp[-\Lambda(t_0, t_a; \theta)].
\end{aligned}$$

A medida que el número de intervalos m aumenta y el tamaño de los intervalos se acerca a cero hay tiempos de recurrencia exactos en $t_1 \leq \dots \leq t_r$ donde $r = \sum_{j=1}^m d_j$, $t_0 \leq t_1, t_1 \leq t_2, \dots, t_r \leq t_a$, entonces utilizando un argumento de limitación se deduce que la probabilidad en términos de la aproximación de la densidad es

$$L(\theta) = \prod_{j=1}^r \lambda(t_j; \theta) \times \exp[-\Lambda(0, t_a; \theta)]$$

2.2.2. Función de verosimilitud del proceso de Poisson superpuesto

Supongamos que se dispone de datos de n sistemas independientes con la misma función de intensidad $\lambda(t)$ y el sistema i se observa en el intervalo $(0, t_a]$, $i = 1, \dots, n$ y los tiempos de recurrencia del sistema i se denotan por $t_{i1}, t_{i2}, \dots, t_{ir}$. Entonces la probabilidad del PPNH es simplemente el producto de las probabilidades de los sistemas individuales

$$L(\theta) = \prod_{i=1}^n \prod_{j=1}^r \lambda(t_{ij}; \theta) \times \exp\left[-\sum_{i=1}^n \Lambda(0, t_{ri}; \theta)\right]$$

La suposición de que todos los sistemas tienen la misma $\lambda(t)$ es una suposición fuerte. Esta suposición a menudo será inapropiada en las aplicaciones prácticas. Las generalizaciones (por ejemplo, el uso de variables explicativas para tener en cuenta las diferencias entre sistemas) son posibles.

2.3. Inferencia en el PP con tasa de fallo potencial

El proceso de ley de potencia es un proceso de Poisson no homogéneo con función de intensidad que tiene la forma

$$\lambda(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}$$

El parámetro β afecta a la forma en que el sistema se deteriora o mejora con el tiempo. Si $\beta > 1$, entonces la función de intensidad $\lambda(t)$ es creciente, y los fallos tienden a producirse con más frecuencia. Si $\beta < 1$, entonces $\lambda(t)$ es decreciente, y el sistema está mejorando. Por último, si $\beta = 1$, el proceso de ley de potencia se reduce al proceso de Poisson homogéneo más simple con intensidad $1/\theta$. El parámetro θ es un parámetro de escala. Hay varias razones por las que el proceso de ley de potencia es un modelo muy utilizado para los sistemas reparables. Duane (1964) demostró que muchos sistemas desarrollados en General Electric parecían seguir un modelo estrechamente relacionado con el proceso de ley de potencia. En lo que se conoció como el gráfico de Duane, sugirió trazar $N(t_i)/t_i$ contra t_i en papel logarítmico doble. Encontró que para muchos sistemas dicho gráfico era casi lineal y sugirió ajustar una línea a los datos. Crow (1974) discutió posteriormente la relación entre el proceso de la ley de potencia y la linealidad en un gráfico de Duane. Otra razón para la popularidad del proceso de ley de potencia es que los procedimientos de inferencia estadística están bien desarrollados y a menudo se pueden utilizar las tablas existentes. También es lo suficientemente flexible para modelar sistemas reparables que se deterioran o mejoran.

2.3.1. Estimación puntual de los parámetros: β y θ

Supongamos que un sistema reparable se observa hasta que ocurren n fallos, por lo que observamos los tiempos de fallo $0 < t_1 < t_2 < \dots < t_n$. El teorema da la distribución de probabilidad conjunta de un PPNH truncado por fallos como

$$f(t_1, t_2, \dots, t_n) = \left(\prod_{i=1}^n \lambda(t_i) \right) \exp \left[- \int_0^{t_n} \lambda(x) dx \right], \quad 0 < t_1 < t_2 < \dots < t_n$$

Para el caso en el que $\lambda(t) = (\beta/\theta)(t/\theta)^{\beta-1}$, tenemos

$$\begin{aligned} f(t_1, t_2, \dots, t_n) &= \left(\prod_{i=1}^n \frac{\beta}{\theta} \left(\frac{t_i}{\theta} \right)^{\beta-1} \right) \exp \left[- \int_0^{t_n} \frac{\beta}{\theta} \left(\frac{x}{\theta} \right)^{\beta-1} dx \right] \\ &= \frac{\beta^n}{\theta^{n\beta}} \left(\prod_{i=1}^n t_i \right)^{\beta-1} \exp \left[- \left(\frac{t_n}{\theta} \right)^\beta \right], \quad 0 < t_1 < t_2 < \dots < t_n \end{aligned}$$

Para obtener los estimadores de máxima verosimilitud, tomamos el logaritmo de esta densidad conjunta y establecemos las primeras derivadas parciales (con respecto a θ y β) iguales a cero. La función de log-verosimilitud es

$$\ell(\theta, \beta | t) = n \log \beta - n\beta \log \theta + (\beta - 1) \sum_{i=1}^n \log t_i - \left(\frac{t_n}{\theta} \right)^\beta$$

Diferenciando con respecto a θ y β , y fijando los resultados iguales a cero, se obtienen las dos ecuaciones

$$0 = \frac{\partial \ell}{\partial \theta} = -\frac{n\beta}{\theta} + \frac{\beta}{\theta} \left(\frac{t_n}{\theta}\right)^\beta$$

y

$$0 = \frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - n \log \theta + \sum_{i=1}^n \log t_i - \left(\frac{t_n}{\theta}\right)^\beta \log \left(\frac{t_n}{\theta}\right)$$

La primera ecuación se simplifica en

$$0 = -n + \left(\frac{t_n}{\theta}\right)^\beta$$

que se puede resolver para θ (en términos de β) para obtener

$$\theta = \frac{t_n}{n^{1/\beta}}$$

Sustituyendo de nuevo en la primera ecuación se obtiene

$$0 = \frac{n}{\beta} - n \log \frac{t_n}{n^{1/\beta}} + \sum_{i=1}^n \log t_i - \left(\frac{t_n n^{1/\beta}}{t_n}\right)^\beta \log \left(\frac{t_n n^{1/\beta}}{t_n}\right)$$

que se simplifica en

$$0 = \frac{n}{\beta} - n \log t_n + \sum_{i=1}^n \log t_i + \frac{n}{\beta} \log n - n \log n^{1/\beta}$$

Los dos últimos términos suman cero. Después de observar que

$$\sum_{i=1}^n \log t_i - n \log t_n = \sum_{i=1}^n \log t_i - \sum_{i=1}^n \log n = \sum_{i=1}^n \log (t_i/t_n)$$

tenemos

$$0 = \frac{n}{\beta} + \sum_{i=1}^n \log (t_i/t_n) = \frac{n}{\beta} - \sum_{i=1}^n \log (t_n/t_i)$$

Resolviendo para β se obtiene

$$\beta = \frac{n}{\sum_{i=1}^n \log (t_n/t_i)}$$

Como esto da β en términos de los tiempos de fallo observados solamente, esto es el estimador de máxima verosimilitud de β . Obsérvese que el último

término de la suma es $\log(t_n/t_n) = \log 1 = 0$, por lo que el estimador de máxima verosimilitud de β puede escribirse como

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n-1} \log(t_n/t_i)}$$

Una vez calculado el estimador de β , se puede volver a sustituir en la expresión para θ , dando el estimador de máxima verosimilitud para θ como

$$\hat{\theta} = \frac{t_n}{n^{1/\hat{\beta}}}$$

Los estimadores de máxima verosimilitud para los parámetros β y θ son por tanto

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \log(t_n/t_i)}$$

y

$$\hat{\theta} = \frac{t_n}{n^{1/\hat{\beta}}}$$

El estimador $\hat{\beta}$ está sesgado. Un estimador insesgado es

$$\bar{\beta} = \frac{n-2}{n} \hat{\beta} = \frac{n-2}{\sum_{i=1}^{n-1} \log(t_n/t_i)}$$

2.3.2. Estimación por intervalos y contrastes de hipótesis

Un resultado útil para derivar intervalos de confianza y pruebas de hipótesis sobre β es que $2n\beta/\hat{\beta}$ tiene una distribución chi-cuadrado con $2(n-1)$ grados de libertad. Para derivar este importante resultado necesitamos el siguiente lema.

Lema 2.1. Si U tiene una distribución uniforme en el intervalo $(0, 1)$, entonces $X = -\theta \log U$ tiene una distribución exponencial con media θ .

Demostración. La función de distribución acumulativa de $X = -\theta \log U$ es

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(-X/\theta \geq -x/\theta) \\ &= P(\exp(-X/\theta) \geq \exp(-x/\theta)) \\ &= P(U \geq \exp(-x/\theta)) \end{aligned}$$

$$= 1 - \exp(-x/\theta), \quad x > 0$$

Esta es la función de distribución acumulativa de la distribución exponencial con media θ .

Teorema 2.2. *Dada una muestra truncada en número de fallos, $2n\beta/\hat{\beta}$ tiene una distribución chi-cuadrado con $2(n-1)$ grados de libertad.*

Demostración. Podemos escribir la expresión $2n\beta/\hat{\beta}$ como

$$\frac{2n\beta}{\hat{\beta}} = 2\beta \sum_{i=1}^n \log(t_n/t_i) = -2\beta \sum_{i=1}^n \log(t_i/t_n)$$

Condicionadas a $T_n = t_n$ las variables aleatorias $T_1 < T_2 < \dots < T_{n-1}$ se distribuyen como estadísticos de orden $n-1$ de la distribución con función de distribución acumulativa

$$G(y) = \begin{cases} 0, & y \leq 0 \\ \Lambda(y)/\Lambda(t_n), & 0 < y < t_n \\ 1, & y \geq t_n \end{cases}$$

Para el proceso de ley de potencia, tenemos $\Lambda(t) = (t/\theta)^\beta$, así que

$$\frac{\Lambda(y)}{\Lambda(t_n)} = \frac{(y/\theta)^\beta}{(t_n/\theta)^\beta} = \left(\frac{y}{t_n}\right)^\beta$$

En este caso tenemos

$$G(y) = \begin{cases} 0 & y \leq 0 \\ (y/t_n)^\beta & 0 < y < t_n \\ 1 & y \geq t_n \end{cases}$$

Siendo Y una variable aleatoria con cdf G , tenemos por un lado

$$\left(\frac{y}{t_n}\right)^\beta = G(y) = P(Y \leq y)$$

para $0 < y < t_n$. Por otra parte,

$$P(Y \leq y) = P(Y/t_n \leq y/t_n) = P\left((Y/t_n)^\beta \leq (y/t_n)^\beta\right)$$

Esto implica que

$$P\left((Y/t_n)^\beta \leq (y/t_n)^\beta\right) = (y/t_n)^\beta$$

para $0 < y < t_n$, o $0 < y/t_n < 1$. Esto significa que la variable aleatoria $(Y/t_n)^\beta$ tiene una distribución uniforme en el intervalo $(0, 1)$. Por lo tanto,

las cantidades $(T_i/t_n)^\beta$, $i = 1, 2, \dots, n-1$ se distribuyen como estadísticas de orden $n-1$ de una distribución uniforme en el intervalo $(0, 1)$. Por el Lema 2.1 antes mencionado, la suma

$$\sum_{i=1}^n -\log(t_i/t_n)^\beta = -\beta \sum_{i=1}^n \log(t_i/t_n)$$

se distribuye, pues, como la suma de $n-1$ variables aleatorias exponenciales, cada una con media 1. Por el Teorema, la suma de $n-1$ variables aleatorias exponenciales tiene una distribución gamma con parámetros $\kappa = n-1$ y $\theta = 1$. Finalmente, el doble de una distribución gamma con parámetros $n-1$ y 1 da lugar a una distribución chi-cuadrado con $2(n-1)$ grados de libertad. Así pues,

$$-2\beta \sum_{i=1}^n \log(t_i/t_n) = \frac{2n\beta}{\hat{\beta}}$$

tiene una distribución $\chi^2(2(n-1))$.

Para obtener un intervalo de confianza de dos colas al nivel $1-\alpha$, buscaremos los valores $\chi_{1-\alpha/2}^2(2(n-1))$ y $\chi_{\alpha/2}^2(2(n-1))$, los puntos que dejan área de $\alpha/2$ en las colas izquierda y derecha, respectivamente, de la distribución chi-cuadrado con $2(n-1)$ grados de libertad. Podemos entonces escribir

$$P\left(\chi_{1-\alpha/2}^2(2(n-1)) < \frac{2n\beta}{\hat{\beta}} < \chi_{\alpha/2}^2(2(n-1))\right) = 1 - \alpha$$

por lo que el intervalo de confianza para β es

$$\frac{\chi_{1-\alpha/2}^2(2(n-1))\hat{\beta}}{2n} < \beta < \frac{\chi_{\alpha/2}^2(2(n-1))\hat{\beta}}{2n}$$

Este intervalo se interpreta como cualquier otro intervalo de confianza. Antes de recoger los datos, la probabilidad de que un intervalo calculado de esta manera incluya el valor verdadero de β es de $1-\alpha$.

Una vez que se recogen los datos, no se puede hacer ninguna declaración de probabilidad porque ya no hay cantidades aleatorias en la expresión del intervalo de confianza. La confianza está en el método que produjo el intervalo, más que en el propio intervalo.

A menudo es interesante ver si el intervalo de confianza para β incluye o excluye el valor $\beta = 1$. Cuando $\beta = 1$ la función de intensidad se reduce a

$$\lambda(t) = (1/\theta)(t/\theta)^{1-1} = 1/\theta$$

que es independiente de t . Así, el proceso de ley de potencia se reduce al proceso de Poisson homogéneo cuando $\beta = 1$.

Podemos contrastar la hipótesis $H_0 : \beta = 1$ contra una de las alternativas $H_1 : \beta \neq 1$, $H_1 : \beta < 1$, o $H_1 : \beta > 1$. (La discusión aquí explica la razón de ser de la prueba MIL-HDBK 189). Si la alternativa es la alternativa de dos lados $H_0 : \beta \neq 1$, entonces como $2n\beta/\hat{\beta}$ tiene una distribución χ^2 con $2(n-1)$ grados de libertad, la regla es rechazar H_0 si

$$2n\beta/\hat{\beta} < \chi_{1-\alpha/2}^2(2(n-1))$$

o si

$$2n\beta/\hat{\beta} > \chi_{\alpha/2}^2(2(n-1))$$

Después de algunas manipulaciones, esto se convierte en

$$\hat{\beta} < \frac{2n}{\chi_{\alpha/2}^2(2(n-1))} \quad \text{o} \quad \hat{\beta} > \frac{2n}{\chi_{1-\alpha/2}^2(2(n-1))}$$

Para probar $H_0 : \beta \leq 1$ frente a $H_1 : \beta > 1$, es decir, una prueba de si el sistema se mantiene estable o posiblemente mejora frente a la alternativa de que se está deteriorando, la regla es rechazar H_0 si

$$\hat{\beta} < \frac{2n}{\chi_{\alpha/2}^2(2(n-1))}$$

El P – *valor* también se puede utilizar para evaluar la fuerza de la evidencia contra la hipótesis nula. Para la alternativa de dos caras, el valor P sería el doble del área de la cola más allá de $2n/\hat{\beta}$. Cuanto menor sea el valor P , más fuerte será la evidencia contra la hipótesis nula.

Los intervalos de confianza y las pruebas de hipótesis para θ pueden construirse a partir del estadístico $W = (\hat{\theta}/\theta)^{hat{\beta}}$. El estadístico W es una cantidad fundamental, es decir, su distribución no depende de ningún parámetro desconocido. Los puntos porcentuales de la distribución de W para varios valores de n fueron determinados por Finkelstein (1976) que utilizó la simulación. Las tablas del artículo de Finkelstein pueden utilizarse para calcular intervalos de confianza para θ . Dichos intervalos de confianza para θ suelen ser extremadamente amplios. Sin embargo, no suele ser necesario estimar θ con gran precisión, ya que este parámetro no tiene una interpretación sencilla.

Tabla: Tiempos de fallo de un generador de avión.

Número de fallo i	Tiempo de fallo t_i	$N(t_i)/t_i$
1	10	0,1000
2	55	0,0364
3	166	0,0181
4	205	0,0195
5	341	0,0147
6	488	0,0123
7	567	0,0123
8	731	0,0109
9	1308	0,0069
10	2050	0,0049
11	2453	0,0045
12	3115	0,0039
13	4017	0,0032
14	4596	0,0030

Ejemplo: A modo de ilustración, considere los tiempos de fallo de un generador de avión que se muestran en la tabla anterior. Estos tiempos de fallo se leyeron de una figura de Duane (1964). La tabla también da los valores de t_i y $N(t_i)/t_i, i = 1, 2, \dots, 14$ para estos datos. Las estimaciones de máxima verosimilitud de β y θ son

$$\hat{\beta} = \frac{14}{\sum_{i=1}^{14-1} \log(4596/t_i)} = \frac{14}{28,97} = 0,483$$

y

$$\hat{\theta} = \frac{4596}{14^{1/0,483}} = \frac{4596}{236,013} = 19,474$$

2.4. Estimación ML para el PPNH con tasa de fallo log-lineal

Supongamos ahora que el PPNH tiner función intensidad log-lineal, es decir

$$\lambda(t) = \exp(\gamma_0 + \gamma_1 t), \quad t > 0.$$

Supongamos que tenemos una muestra de tiempos de ocurrencia $t_1 < t_2 < \dots < t_n$, que han sido generados desde este modelo. Con un $\lambda(t)$ log-lineal y tiempos de recurrencia exactos, la verosimilitud asociada a los tiempos de recurrencia registrados para un solo sistema es

2.5. TRATAMIENTO COMPUTACIONAL: IMPLEMENTACIÓN EN R37

$$L(\gamma_0, \gamma_1) = \exp\left(r\gamma_0 + \gamma_1 \sum_{j=1}^r t_j\right) \times \exp[-\Lambda(t_a; \gamma_0, \gamma_1)].$$

Las estimaciones ML $\hat{\gamma}_0$ y $\hat{\gamma}_1$ se obtienen resolviendo

$$\sum_{j=1}^r t_j + \frac{r}{\gamma_1} - \frac{rt_a \exp(\gamma_1 t_a)}{\exp(\gamma_1 t_a) - 1} = 0$$

y

$$\gamma_0 = \log\left(\frac{r\gamma_1}{\exp(t_a \gamma_1) - 1}\right)$$

La probabilidad relativa es:

$$R(\gamma_0, \gamma_1) = \exp\left[r(\gamma_0 - \hat{\gamma}_0) + (\gamma_1 - \hat{\gamma}_1) \sum_{j=1}^r t_j\right] \times \exp\{r - \Lambda(t_a; \gamma_0, \gamma_1)\}.$$

Una predicción puntual del número futuro de recurrencias utilizando el modelo loglineal PPNH es

$$\int_a^b \lambda(u, \hat{\theta}) du = \frac{\exp(\hat{\gamma}_0)}{\hat{\gamma}_1} [\exp(\hat{\gamma}_1 b) - \exp(\hat{\gamma}_1 a)]$$

2.5. Tratamiento computacional: implementación en R

En esta sección presentamos un código en R que permite obtener los estimadores presentados en la sección anterior así como llevar a cabo algunas inferencias basadas en estos estimadores. En concreto nos centraremos en el modelo con función intensidad potencial.

```
1  ### Ajuste parametrico de la funcion ROCOF:
2  ### Modelo de tendencia potencial: lambda(t)= (theta1/
   theta2)*(t/theta2)^(theta1-1)
3  ajuste.pot<-function(data, aggregated=F)
4  {
5     if(aggregated==F) ## data es un vector con los tiempos
   de ocurrencias
6     {ti<-sort(data);n<-length(ti)
7     t.max<-ti[n]
```

```

8
9  theta1<-n/(sum(log(t.max/ti)))
10 theta2<-t.max/(n^(1/theta1))
11 theta<-round(c(theta1=theta1,theta2=theta2),3)
12 Lambda.i<- (ti/theta2)^theta1
13 x11()
14 plot(c(0,ti),0:n,type='s',xlab='t',ylab='N(t)',
15      lty=1,lwd=2,main='Ajuste potencial a la funcion
      ROCOF acumulada (tiempos de suceso registrados
      )')
16
17 lines(c(0,ti),c(0,Lambda.i),lwd=2,lty=2,col=2)
18 legend('topleft',c(paste('Ajuste por ML: theta1=',
19      theta[1],'; theta2=',theta[2],sep=''),
20      'Curva empirica'),lty=c(2,1),
21      lwd=c(2,2),col=c(2,1),bty=
22      'n')
23
24 return(theta=theta)
25 }else{ ##aggregated=T --> Los datos son una matriz n x
26      2: datos=(Ei,ni)
27
28      ## Ei= los extremos de los
29      intervalos;
30      ## ni= numero de ocurrencias
31      en cada intervalo
32
33      ti<-c(0,data[,1]);di<-data[,2];n<-length(ti)
34      ## funcion de verosimilitud
35
36      loglik<-function(pars){
37      theta1<-pars[1];theta2<-pars[2]
38      Lambda<-function(t){(t/theta2)^theta1}
39      ll<-sum(di*log(Lambda(ti[-1])-Lambda(ti[-n])),na.rm=
40      T)-Lambda(ti[n])
41      return(-ll)
42      }
43
44      library(Rsolnp) ##para obtener el \'optimo
45      pars<-c(1,1)
46      res<-solnp(pars=pars, fun = loglik, LB =c(0,0),
47      control=list(trace=0))
48      theta<-round(res$pars,3) ##parametros
49      estimados: theta1 y theta2
50      Lambda.i<- (ti/theta[2])^theta[1] ## la funcion
51      intensidad acumulada
52
53      x11()
54      ## Curva empirica:

```

```

39     plot(c(ti),c(0,cumsum(di)),col=1,lty=1,lwd=2,type='s
      ',xlab='t',ylab='N(t)',
40         main='Ajuste potencial a la funcion ROCOF
            acumulada (datos agregados)')
41     ## Curva ajustada:
42     lines(c(ti),c(Lambda.i),lwd=2,lty=2,col=2)
43     legend('topleft',c(paste('Ajuste por ML: theta1=',
            theta[1],'; theta2=',theta[2],sep='')),
44           'Curva empirica'),lty=c(2,1),lwd
            =c(2,2),col=c(2,1),bty='n')
45     return(theta=theta)
46 }
47 }

```

2.6. Pruebas de la tendencia de la tasa de fallo

El gráfico más sencillo para los datos de recurrencia de un solo sistema muestra el número acumulado de recurrencias del sistema en función del tiempo (un caso especial de una función acumulativa media MCF utilizado con múltiples sistemas). La no linealidad de este gráfico indica que los tiempos de interrecurrencia no están idénticamente distribuidos. Para los procesos de Poisson, esto implica una tasa de recurrencia no constante. Un PPH debe tener un gráfico MCF que sea aproximadamente lineal (pero una MCF lineal por sí sola no implica un PPH). Un gráfico de tiempos de interrecurrencia frente a la edad del sistema o un "gráfico de series temporales" que muestre los tiempos de interrecurrencia frente al número de recurrencias permitirá descubrir tendencias o ciclos que sugieran que los tiempos de interrecurrencia no están distribuidos de forma idéntica.

Existen varias pruebas formales de tendencia.

2.6.1. El test *Military Handbook*

Cuando se elige un modelo de ley de potencia del PPNH o sin tendencia, la prueba de Military Handbook es mejor para encontrar significancia (Duane). En otras palabras, si los datos provienen de un sistema que sigue la ley de potencia, esta prueba generalmente funcionará mejor que cualquier otra prueba en términos de encontrar significancia.

La prueba *Military Handbook* comprueba para $\beta = 1$ (implica un PPH, y por lo tanto no hay tendencia) en el modelo PPNH de potencia. La estadís-

tica $X_{MHB}^2 = -2 \sum_{j=1}^r \log(t_j/t_a)$ tiene aproximadamente una distribución $\chi_{(2r)}^2$ bajo el modelo PPH. Así los valores de X_{MHB}^2 mayores que $\chi_{(1-\alpha;2r)}^2$ proporcionan evidencia de una tasa de recurrencia no constante $\lambda(t)$ al nivel de significancia del $100\alpha\%$. Esta es una poderosa herramienta para probar el PPH frente a la PPNH con un modelo de potencia $\lambda(t)$.

2.6.2. El test de Laplace

Esta prueba es mejor para encontrar significancia cuando la elección es entre sin tendencia y un modelo exponencial PPNH. Es decir que, si los datos provienen de un sistema que sigue la Ley Exponencial, esta prueba generalmente funcionará mejor que cualquier prueba en términos de encontrar significancia.

La conocida prueba de Laplace tiene una base similar para comprobar la tendencia en el modelo loglineal PPNH. En este caso, si el proceso subyacente es PPH ($\gamma_1 = 0$), la estadística de prueba $Z_{LP} = \frac{\sum_{j=1}^r t_j/t_a - r/2}{\sqrt{r/12}}$ tiene, aproximadamente, una distribución $NOR(0, 1)$. Así, los valores de Z_{LP} superiores a $z_{(1-\alpha/2)}$ proporciona pruebas de una $\lambda(t)$ no constante a un nivel de significancia al $100\alpha\%$. Esta es una herramienta poderosa para probar el PPH frente al PPNH con una $\lambda(r)$ loglineal.

Tanto la prueba del *Military Handbook* como la prueba de Laplace pueden dar lugar a conclusiones erróneas en situaciones en las que no hay tendencia pero el proceso subyacente es un proceso de renovación distinto de la del PPH.

2.6.3. El test de Lewis-Robinson

La prueba de Lewis-Robinson para la tendencia utiliza $Z_{LR} = Z_{LP} \times \frac{\bar{\tau}}{S_{\tau}}$, donde $\bar{\tau}$ y S_{τ} son, respectivamente, la media muestral y la desviación estándar de los tiempos de interrecurrencia. En muestras grandes, Z_{LR} sigue aproximadamente una distribución $NOR(0, 1)$ (norma estándar) si el proceso subyacente es un proceso de renovación (los procesos de renovación, en general, no tienen tendencia). El estadístico Z_{LR} se obtuvo a partir de argumentos heurísticos para tener en cuenta los tiempos de interrecurrencia no exponenciales, ajustándolos a un coeficiente de variación diferente (la distribución exponencial tiene un coeficiente de variación igual a 1). Los resultados de Lawless y Thiagarajah (1996) indican que Z_{LR} es preferible a Z_{LP} como prueba general de tendencia en datos de procesos puntuales.

2.7. Prueba de tiempos de interurrencia independientes

Al evaluar la idoneidad de un modelo de proceso de renovación, también es necesario comprobar si la hipótesis del modelo de tiempos de interurrencia independientes es coherente con los datos. Para ello, se considera la correlación serial en la secuencia de tiempos de interurrencia. El trazado de los tiempos de interurrencia frente a los tiempos de interurrencia retardados (τ_i versus τ_{i+k}) proporciona una representación gráfica de la correlación en serie (la correlación entre tiempos de interurrencia adyacentes). El coeficiente de correlación serial de $lag - k$ es $\rho_k = \frac{Cov(\tau_j, \tau_{j+k})}{Var(\tau_j)}$. La correlación serial de primer orden ($k = 1$) suele ser el retardo más importante a considerar. Si $\tau_1, \tau_2, \dots, \tau_r$ son los tiempos de interurrencia observados, el coeficiente de correlación serial de la muestra es $\hat{\rho}_k = \frac{\sum_{j=1}^{r-k} (\tau_j - \bar{\tau})(\tau_{j+k} - \bar{\tau})}{\sqrt{\sum_{j=1}^{r-k} (\tau_j - \bar{\tau})^2 \sum_{j=1}^{r-k} (\tau_{j+k} - \bar{\tau})^2}}$, donde $\bar{\tau} = \sum_{j=1}^r \frac{\tau_j}{r}$. Cuando $\rho_k = 0$ y r es grande, $\sqrt{r-k} \times \hat{\rho}_k \sim NOR(0, 1)$. Esta distribución aproximada puede utilizarse para evaluar si ρ_k es diferente de cero.

2.8. Generación de trayectorias muestrales de un PPNH

Las realizaciones pseudoaleatorias (pseudorandom) son útiles para comprobar la adecuación de las aproximaciones de grandes muestras y para implementar métodos bootstrap.

Utilizando las propiedades del modelo PPNH, se puede demostrar que para una $\Lambda(t)$ monótona creciente, las variables aleatorias $\Lambda(T_{i-1}, T_i)$, $i = 1, 2, \dots$, son iid cada una con $EXP(1)$ donde $T_0 = 0$. Supongamos que U_i , $i = 1, \dots, r$ es una muestra pseudoaleatoria de un $UNIF(0, 1)$. Entonces utilizando $\Lambda(T_{i-1}, T_i) = \Lambda(T_i) - \Lambda(T_{i-1})$ y resolviendo secuencialmente para T_1, T_2, \dots se obtienen las ecuaciones

$$\begin{aligned}\Lambda(T_1) &= -\log(U_1) \\ \Lambda(T_2) - \Lambda(T_1) &= -\log(U_2) \\ &\vdots \\ \Lambda(T_r) - \Lambda(T_{r-1}) &= -\log(U_r)\end{aligned}$$

Entonces, resolviendo para T_1, \dots, T_r se obtienen r tiempos de recurrencia del PPNH con la tasa de recurrencia $\lambda(t)$ o $(MCF\Lambda(t))$. Si se desea una realización en un intervalo $(0, t_a]$, entonces r es aleatoria y el proceso secuencial se detiene cuando $T_i > t_a$, lo que implica que el número de recurrencias es $r = i - 1$. La solución general puede expresarse como

$$T_j = \Lambda^{-1} \left[- \sum_{i=1}^j \log(U_i) \right], j = 1, \dots$$

Se puede expresar la solución general recursivamente como

$$T_j = \Lambda^{-1} [\Lambda(T_{j-1}) - \log(U_j)], j = 1, \dots$$

donde $T_0 = 0$.

2.8.1. PPNH con una tasa de fallo potencial

En este caso $\Lambda(t) = \left(\frac{t}{\eta}\right)^\beta$ y $\Lambda^{-1}(s) = \eta \times s^{(1/\beta)}$. Entonces utilizando la solución general se obtiene

$$T_j = \eta \times \left[- \sum_{i=1}^j \log(U_i) \right]^{1/\beta} = \left[T_{j-1}^\beta - \eta^\beta \times \log(U_j) \right]^{1/\beta}, j = 1, \dots$$

2.8.2. PPNH con una tasa de fallo log-lineal

En este caso

$$\Lambda(t) = [\exp(\gamma_0)] [\exp(\gamma_1 t) - 1] / \gamma_1$$

y

$$\Lambda^{-1}(s) = (1/\gamma_1) \log [\gamma_1 \exp(-\gamma_0) s + 1]$$

Utilizando de nuevo la solución general, se obtiene

$$T_j = \frac{1}{\gamma_1} \times \log \left[-\gamma_1 \times \exp(-\gamma_0) \times \sum_{i=1}^j \log(U_i) + 1 \right]$$

$$= \frac{1}{\gamma_1} \times \log [\exp(\gamma_1 T_{j-1}) - \gamma_1 \times \exp(-\gamma_0) \times \log(U_j)], j = 1, \dots$$

2.8. GENERACIÓN DE TRAYECTORIAS MUESTRALES DE UN PPNH43

2.8.3. Generación de trayectorias muestrales de procesos de Poisson en R

Aquí las funciones de R que hemos hecho para generar los datos y algunos ejemplos.

```
1 #rm(list=ls(all=TRUE))
2 NHPP.sim<-function(n,model,theta1,theta2,pinta='si',
3   aggregated=F)
4 {
5   ## esta funci\on simula trayectorias de un proceso de
6   Poisson no homogeneo
7   ## n es el tamano de la trayectoria (muestra)
8   ## consideramos 3 modelos para la funcion intensidad,
9   lambda(t), es decir el ROCOF, m(t)
10  ## que en el proceso de Poisson son la misma funcion.
11  ## modelo potencial, modelo log-lineal y un modelo
12  periodico
13  ## en los dos primeros el usuario puede elegir los
14  valores de los parametros: theta1 y theta2
15  ## en el modelo periodico no.
16
17  if (model=='pot') #m(t)=lambda(t)=(theta1/theta2)*(t/
18  theta2)^(theta1-1)
19  #M(t)=Lambda(t)=(t/theta2)^theta1
20  {lambda<-function(t) {(theta1/theta2)*(t/theta2)^(
21  theta1-1)} #funcion intensidad del proceso -->
22  rocof
23  Lambda<-function(t) {(t/theta2)^theta1} #funcion
24  intensidad acumulada del proceso
25
26  Lambda.inv<-function(y) (y^(1/theta1))*theta2 #
27  transformacion inversa --> rocof
28  } ## este trozo de codigo espesifica la funcion
29  intensidad lambda
30  ## la funcion intensidad acumulada Lambda=int_0^t
31  lambda
32  ## y la funcion inversa de la funcion Lambda:
33  Lambda.inv(Lambda(t))=t
34  ## Hacemos lo mismo para los dos modelos restantes.
35  if (model=='log-lin')
36  {
37    lambda<-function(t) {exp(theta1+theta2*t)} #funcion
```

```

    intensidad del proceso
25   Lambda<-function(t) {(exp(theta1)/theta2)*(exp(
      theta2*t)-1)}
26                                     # {exp(t)-1} #funcion
                                       intensidad acumulada del
                                       proceso
27
28   Lambda.inv<-function(y) {((log(exp(theta1)+theta2*y)
      )-theta1)/theta2}
29                                     # {log(y+1)} #transformacion
                                       inversa
30 } ## theta1=0;theta2=1
31
32 if(model=='periodico') ## este modelo solo
    consideramos un modelo,
33                                     ## no cambiamos los parametros
34 {
35   theta1=NA;theta2=NA
36   lambda<-function(t) cos(t/2)+2 #funcion intensidad
    del proceso
37   Lambda<-function(t) 2*t+2*sin(t/2) #funcion
    intensidad acumulada del proceso
38
39   Lambda.inv<-function(y) #transformacion inversa
    {f.aux<-function(t)return(2*t+sin(t/2)*2-y)
40   s.y<-uniroot(f.aux,lower=0,upper=1000)$root
41   return(s.y)
42 }
43 }
44 }
45 #genero n tiempos de un proceso PPH(1)
46
47 ## Una secuencia de variables Exp(1) independientes
48 ## determina los tiempos entre ocurrencias de un PP(1)
49 Xn<-rexp(n,rate=1) #exponenciales iguales
    razon=1
50
51 ## Y.n=X.1+X.2+...+X.n es el tiempo de ocurrencia de
    la n-esima llegada del proceso PP(1)
52 Yn<-cumsum(Xn) #tiempos de llegadas del PPH(1)
53
54 ## Propiedad para generar la trayectoria del proceso
    PPNH de razon lambda(t):
55 ## T1< T2 <... <Tn son los tiempos de ocurrencia de un

```

2.8. GENERACIÓN DE TRAYECTORIAS MUESTRALES DE UN PPNH45

```
PPNH(lambda(t))
56  ## entonces Lambda(T1)=Y1 < Lambda(T2)=Y2 < ...<
    Lambda(Tn)=Yn son los tiempos de ocurrencia
57  ## de un proceso PP(1)
58  ## Arriba hemos generado la secuencia Yn; si calculo
    la inversa mediante Lambda.inv de Yn
59  ## tendremos los tiempos del proceso deseado T1 < T2<
    ... <Tn
60  #
61  ## T.n<-sapply(Y.n,Lambda.inv)
62
63  ### otra forma
64  Tn<-double(n)
65  for (i in 1:n) {Tn[i]<-Lambda.inv(Yn[i])}
66
67
68  #comprobacion
69  if (pinta=='si')
70  {## pinto primero la trayectoria empirica del proceso,
    es decir los datos:
71    plot(c(0,Tn),0:n,type='s',xlab='t',lwd=2,ylab='N(t)'
    ,
72    main=paste('Modelo: ', model,sep=''),lty=2,col
    =2)
73    ## a continuacion la curva teorica:
74    lines(c(0,Tn),c(0,Lambda(Tn)),lwd=2,lty=1,col=1)
75    ### las dos curvas deben estar muy proximas entre si.
76    ## cuanto mayor sea el valor de n (tamano de la
    muestra), mas cerca.
77    legend('topleft',c('ROCOF acumulado teorico',
78    'ROCOF acumulado empirico',
79    paste('Parametros: (theta1=',theta1
    ,', theta2=',theta2,')',sep=''))
    ,
80    lwd=c(2,2),lty=c(1,2,NA),col=c(1,2,NA),
    bty='n')
81  }
82  #resultado
83  ## la funcion tambien devuelve los valores de la
    trayectoria muestral generada
84  if(aggregated==F){return(Tn)} ### devolvemos los
    tiempos de ocurrencia ordenados o bien
85  if(aggregated==T){
```

```

86   ti<-seq(0,max(Tn),length=101)
87   di<-hist(x=Tn,breaks=ti,right=T,plot=F)$counts
88   recuentos<-data.frame(ti[-1],c(di));
89   colnames(recuentos)<-c('tiempos','n.eventos')
90   return(recuentos)
91   }
92 }

```

2.8.4. Ejemplo práctico

Se simulan los procesos de Poisson no homogéneos, para una muestra de 300 datos, con el modelo potencial (tasa de fallo potencial), $\theta_1 = 0,8$ y $\theta_2 = 1$.

Seguidamente se presenta la Figura 2.8.4. Se consideran 3 modelos distintos desde los que generamos muestras de tamaño 300 en cada caso. Se genera el ROCOF acumulado empírico o tasa de fallos acumulada, es decir el número acumulado de ocurrencias por unidad de tiempo, lo que proporciona, como puede verse en la Figura 2.8.4, curvas escalonadas. Observamos también en todos los casos que la curva se acerca razonablemente bien a la curva del ROCOF acumulado teórico, con algunas discrepancias en determinadas zonas. En el modelo potencial el parámetro $\theta_1 = 0,8$ de la leyenda del gráfico se corresponde con el parámetro de forma, mientras que el valor de $\theta_2 = 1$ se corresponde con el parámetro de escala de modo que el ROCOF o tasa de fallos teórica es una función que presenta una tendencia decreciente en la ocurrencia de fallos con el tiempo. Específicamente en este caso consideramos una tasa de ocurrencia de fallos de la forma $\lambda_1(t) = 0,8t^{-0,2}$. Para el modelo log-lineal, la elección de los parámetros $\theta_1 = 0$ $\theta_2 = 1$ se corresponden con un modelo de tasa de fallo (ROCOF) $\lambda_2(t) = \exp(t)$. Por último, el modelo que denominamos periódico está determinado por una tasa de fallo $\lambda_3(t) = \cos(t/2) + 2$.

Se hace el ajuste paramétrico al modelo potencial, teniendo en cuenta que contamos con los tiempos exactos de ocurrencia de fallos, obtenemos como resultados los valores estimados para los parámetros $\theta_1 = 0,848$ y $\theta_2 = 1,464$. Se compara este ajuste con el modelo empírico $\Lambda(t)$, es decir, tomamos la función intensidad acumulada con los parámetros ajustados y la función intensidad acumulada empírica. Veamos que, el ajuste paramétrico y el ROCOF acumulado que se obtuvo antes, parecen alinearse bien entre ellos.

Se calculan los Test de tendencias: Laplace, Military Handbook y Lewis-Robinson, respectivamente, implementados tal como hemos explicado en sec-

2.8. GENERACIÓN DE TRAYECTORIAS MUESTRALES DE UN PPNH47

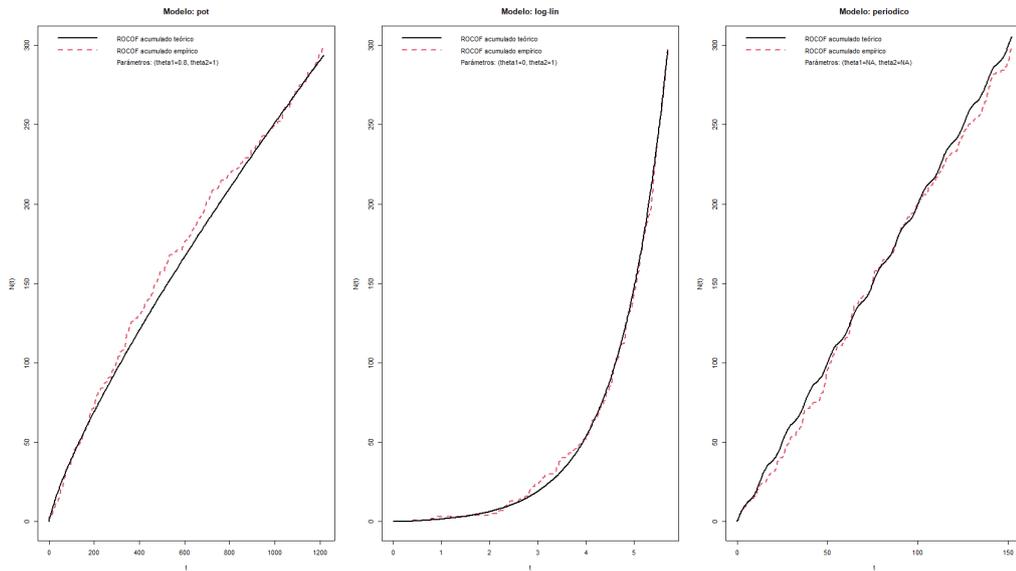


Figura 2.1: Panel izquierdo: modelo potencial. Panel central: modelo log-lineal. Panel derecho: modelo periódico. Todos los paneles: la curva negra continua representa el ROCOF acumulado teórico y la curva roja escalonada es la función ROCOF acumulada empírica.

ciones anteriores. Hemos considerado los datos generados desde el modelo potencial con tasa de fallo decreciente, es decir $\lambda_1(t)$. Como resultado se tiene que para el Test de Laplace el *pvalor* es 0,002082703 por lo que se rechaza que haya una tasa de recurrencia constante, como se indica. También con el test Military Handbook y el test Lewis-Robinson, los *pvalores* son 0,001331542 y $3,313762e - 06$, con lo cual se afirma que la tasa de fallos no es constante para estos datos.

```
> LP
$Test.value
[1] -3.078178

$p.value
[1] 0.002082703

$Decision
[1] "Reject constant trend"

> MHB<-test.MHB(data=tiempos,alpha=0.05)
```

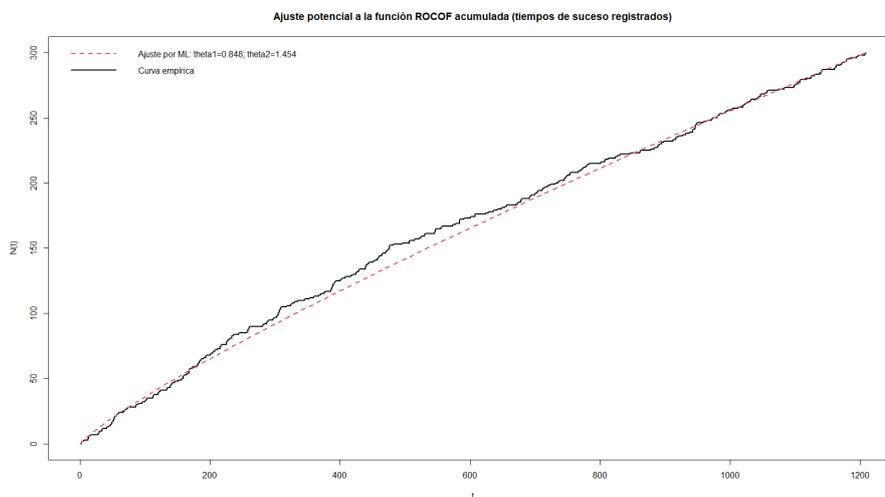


Figura 2.2: Ajuste paramétrico basado en el modelo potencial

```

> MHB
$Test.value
[1] 707.2611

$p.value
[1] 0.001331542

$Decision
[1] "Reject constant trend"

> LR<-test.LR(data=tiempos,alpha=0.05)
> LR
$Test.value
[1] -4.650348

$p.value
[1] 3.313762e-06

$Decision
[1] "Reject constant trend"

```

Capítulo 3

Inferencia no paramétrica en el PPNH

3.1. Estimación suave de la función ROCOF

Para lograr la estimación no paramétrica de la ROCOF se considera una o varias realizaciones de un PPNH, lo que significa observar un único sistema o, alternativamente, una población de sistemas con las mismas características. Por otro lado, no se adopta ningún supuesto para la forma funcional de la función ROCOF $\lambda(t)$, excepto que es una función suave, definiéndose la suavidad en términos de algunas propiedades de derivabilidad de la función con respecto a t .

3.1.1. Descripción del modelo

Se consideran las observaciones del tiempo de fallo de los sistemas. Los tiempos de reparación se estiman depreciables, por lo que no son incluidos en el análisis de la fiabilidad del sistema. Se supone que los fallos que se producen en intervalos disjuntos son independientes. En el sistema no hay fallos simultáneos, y una vez que el sistema falla y luego vuelve a funcionar, su estado es el mismo que tenía antes del fallo. El entorno del sistema adopta la condición de "tan malo como el anterior" tras el mantenimiento de fallo y reparación o, lo que es lo mismo, de reparación mínima. Observemos un sistema durante un periodo de prueba de $(0, \tau]$ en el que se ha registrado un número total de n fallos en tiempos $S_1 < S_2 < \dots < S_n < \tau$. Así, la información de muestreo consistirá en una única realización de un PPNH, con el objetivo de obtener una estimación de la función ROCOF $\lambda(t)$.

En el caso de que los datos estén regularmente espaciados, es decir, que los fallos se produzcan regularmente en el tiempo, se puede construir un

estimador de la ROCOF utilizando funciones *kernel*, es decir, definir

$$\hat{\lambda}_h(t) = \frac{1}{h} \sum_{i=1}^n K\left(\frac{S_i - t}{h}\right),$$

donde K es la función de densidad acotada, no negativa, inicialmente simétrica, y h es un parámetro de banda que se encarga de cuantificar la suavidad de $\hat{\lambda}_h(t)$. Hay que señalar que el aspecto del estimador, es similar al típico estimador de densidad de núcleo, excepto por el factor de normalización que no es necesario al definir $\hat{\lambda}_h(t)$, ya que no se espera que la función λ integre 1, al contrario, $\int_0^\tau \lambda(u)du = E[N(\tau)]$. Por otra parte, el tamaño de la muestra, es decir, el número total de fallos observados en el intervalo fijo $[0, \tau]$ es aleatorio.

El método del kernel aproxima el valor de la ROCOF en cada punto de tiempo t utilizando una función constante. Las estimaciones se calculan localmente y el parámetro de ancho de banda h controla el tamaño de la vecindad local, mientras que el kernel K asigna un peso a cada tiempo de fallo observado. Cuando los fallos se producen regularmente a lo largo del periodo de prueba $(0, \tau]$, el estimador kernel proporciona una aproximación satisfactoria de la ROCOF, Gámiz *et al.* (2011).

En situaciones prácticas es muy común que existan periodos de tiempo concretos en la vida del sistema en los que los fallos se producen con mayor frecuencia. En consecuencia, la muestra se compone de datos que resultan ser muy densos en algunos períodos, mientras que se vuelven más escasos en otros periodos de tiempo. En Gámiz *et al.* (2011) consideran una transformación logarítmica de la función de intensidad, es decir, se define $\varphi(t) = \ln \lambda(t)$. La función λ se supone que es suave en el sentido de ciertas propiedades de derivabilidad, en particular se supone que la derivada de primer orden existe. Por lo tanto, la función λ puede ser aproximada, mediante la expansión de Taylor, por una función lineal de la forma $\varphi(t) \approx \alpha + \beta(t - t_0)$, y $\hat{\lambda}(t_0) = \exp(\hat{\alpha})$.

Dado un conjunto de datos consistente en los tiempos de espera de un PPNH (los tiempos de fallo sucesivos de un sistema en reparación mínima) registrados en un intervalo de prueba como $(0, \tau]$, es decir $0 = S_0 < S_1 < \dots < S_n < S_{n+1} = \tau$, la función de verosimilitud puede expresarse como

$$L = \prod_{i=1}^n \lambda(S_i) \times \exp \left[- \sum_{i=0}^n \int_{S_i}^{S_{i+1}} \lambda(u)du \right] = \prod_{i=1}^n \lambda(S_i) \times \exp \left[- \int_0^\tau \lambda(u)du \right],$$

se puede expresar la función de log-verosimilitud en términos de $\varphi(t)$, esto es,

$$\ell = \ln L = - \int_0^\tau e^{\varphi(u)} du + \sum_{i=1}^n \varphi(S_i).$$

Se puede construir una versión local de la verosimilitud introduciendo una función de peso no negativa adecuada K y un parámetro de ancho de banda h , es decir que, se puede obtener la función de verosimilitud logarítmica ponderada local

$$\ell_h(t_0) = - \int_0^\tau \exp[\alpha + \beta(u - t_0)] K_h(u - t_0) du + \sum_{i=1}^n [\alpha + \beta(S_i - t_0)] K_h(S_i - t_0),$$

donde $K_h(\cdot) = K(\cdot/h)/h$ es una función de núcleo que determina los pesos asignados a las observaciones en torno a t_0 y, en cierto sentido, garantiza que la aproximación lineal de la función γ se aplica en la vecindad donde es válida.

3.1.2. Implementación en R

Esta sección incluye un código en el que implementamos la estimación no paramétrica de la función ROCOF explicada anteriormente considerando un ajuste localmente constante.

```

1  ### 1. Estimacion nucleo de la funcion ROCOF. Un sistema
2  lambda.t<-function(t,datos,h){
3    K<-function(x) return(dnorm(x,0,1))
4    m<-length(t)
5    lambda<-sapply(1:m,function(i) {(1/h)*sum(K((t[i]-
6      datos)/h))})
7    return(lambda)
8  }
9
10 ## estimacion tipo nucleo de la funcion intensidad
11   acumulada: Lambda(t)=int_0^t lambda(s) ds
12 Lambda.t<-function(t,datos,h)
13 {
14   # t es el punto en el que estimamos la curva
15   # ti son los datos, es decir los tiempos de ocurrencia
16   # de fallos
17   # h es el ancho de banda, controla la suavidad de la
18   # curva estimada
19   K.dist<-function(x){pnorm(x,0,1,lower.tail=T)}
20   m<-length(t)

```

```

18     Lambda<-sapply(1:m,function(i) {sum(K.dist((t[i]-datos
                  )/h))})
19     return(Lambda)
20     }

```

3.2. El método de validación cruzada

En esta sección describimos el método de validación cruzada para la estimación del parámetro ancho de banda.

3.2.1. Descripción del modelo

Para estimar el parámetro de suavizado, se considera el método de validación cruzada por mínimos cuadrados que fue introducido por Rudemo (1982) y Bowman (1984). El método es un procedimiento completamente automático (o basado en datos) para elegir un valor adecuado para el parámetro de ancho de banda. Brooks y Marron (1991) adoptan como medida del error de estimación el Error Cuadrado Integrado (ISE), que se define en el contexto de las funciones de intensidad como:

$$ISE(h) = \int_0^\tau (\hat{\lambda}_h(t) - \lambda(t))^2 dt$$

Así, bajo condiciones suaves en el núcleo, muestran con probabilidad uno,

$$\frac{ISE(h_{cv})}{ISE(h_{opt})} \rightarrow 1, \text{ cuando } n \rightarrow \infty$$

donde (h_{cv}) es el valor del ancho de banda obtenido mediante la técnica de validación cruzada, y h_{opt} es el valor óptimo del ancho de banda, es decir, el que que minimiza el ISE . Los fundamentos de la técnica de validación cruzada por mínimos cuadrados adaptada a la estimación del kernel de intensidad son: el ISE se puede escribir

$$ISE(h) = \int_0^\tau \hat{\lambda}_h(t)^2 dt - 2 \int_0^\tau \hat{\lambda}_h(t)\lambda(t)dt + \int_0^\tau \lambda(t)^2 dt$$

donde $\int_0^\tau \lambda(t)^2 dt$ no depende de h por lo que una elección ideal del parámetro de ancho de banda (en el sentido de minimizar el ISE) correspondería a la elección que minimiza los dos primeros términos

$$\int_0^\tau \hat{\lambda}_h(t)^2 dt - 2 \int_0^\tau \hat{\lambda}_h(t)\lambda(t)dt$$

El principio básico del método de validación cruzada por mínimos cuadrados consiste en obtener de los datos una estimación del término $-2 \int_0^\tau \hat{\lambda}_h(t) \lambda(t) dt$ sustituir el término correspondiente y luego minimizar la expresión resultante sobre h . Teniendo esto en cuenta, se define $\hat{\lambda}_{-i}$ como la estimación de la intensidad construida a partir de todos los puntos de datos excepto S_i , es decir, el estimador "leave-one-out" dado por

$$\hat{\lambda}_{-i}(t) = \frac{1}{h} \sum_{j \neq i}^n K \left(\frac{t - S_j}{h} \right)$$

Se define la función de puntuación de validación cruzada sugerida en Brooks y Marron (1991), esta es

$$CV(h) = \int_0^\tau \hat{\lambda}_h(t)^2 dt - 2 \sum_{i=1}^n \hat{\lambda}_{-i}(S_i)$$

donde la función de puntuación depende únicamente de los datos. Además, se puede comprobar que

$$E \left(\sum_{i=1}^n \hat{\lambda}_{-i}(S_i) \right) = E \left(\int_0^\tau \hat{\lambda}_h(t) \lambda(t) dt \right)$$

y luego, sustituyendo la expresión anterior en la definición de $CV(h)$ se observa que $CV(h) + \int_0^\tau \lambda^2(t) dt$ es una buena aproximación del error cuadrático medio integrado ($MISE$)

$$MISE(h) = E \left[\int_0^\tau (\hat{\lambda}_h(t) - \lambda(t))^2 dt \right]$$

Finalmente, suponiendo que el minimizador de $E[CV(h)]$ está cerca del minimizador de $CV(h)$ se puede esperar que este último proporcione una buena elección para el parámetro de ancho de banda.

En Gámiz *et al.* (2011) suponen que la función *kernel* K es simétrica. Haciendo $u = \frac{t}{h}$ el término $\int_0^\tau \hat{\lambda}_h(t)^2 dt$ se tiene lo siguiente

$$\begin{aligned} \int_0^\tau \hat{\lambda}_h(t)^2 dt &= \int_0^\tau \sum_{i=1}^n h^{-1} K \left(\frac{t - S_i}{h} \right) \times \sum_{j=1}^n h^{-1} K \left(\frac{t - S_j}{h} \right) dt \\ &= h^{-1} \sum_{i=1}^n \sum_{j=1}^n \int_0^\tau K \left(h^{-1} S_i - u \right) K \left(u - h^{-1} S_j \right) du \\ &= h^{-1} \sum_{i=1}^n \sum_{j=1}^n K^{(2)} \left(\frac{S_j - S_i}{h} \right) \end{aligned}$$

donde $K^{(2)}$ es la convolución del núcleo consigo mismo. Entonces se puede expresar

$$\begin{aligned}\sum_{i=1}^n \hat{\lambda}_{-i}(S_i) &= \sum_{i=1}^n \left\{ \sum_{j=1}^n h^{-1} K\left(\frac{S_j - S_i}{h}\right) - h^{-1} K(0) \right\} \\ &= \sum_{i=1}^n \sum_{j=1}^n h^{-1} K\left(\frac{S_j - S_i}{h}\right) - nh^{-1} K(0)\end{aligned}$$

y, por último,

$$CV(h) = h^{-1} \sum_{i=1}^n \sum_{j=1}^n \left\{ K^{(2)} - 2w \right\} \left(\frac{S_j - S_i}{h} \right) + 2nh^{-1} K(0)$$

3.2.2. Implementación en R

El código en lenguaje R que presentamos a continuación implementa el método de validación cruzada explicado en la sección anterior.

```

1  ## Selección del ancho de banda por el metodo de
   validacion cruzada
2  ## Necesitamos un valor de h para la funcion 'lambda.t'
   del apartado anterior
3  h.cv<-function(datos){
4    CV<-function(h){
5      cv1.d<-NA
6      lambda.dat<-sapply(datos, function(x) return(lambda.t
7        (x,datos,h)))
8      cv1.d<-sum((lambda.dat^2)*diff(c(0,datos)),na.rm=
9        TRUE)
10     if (is.na(cv1.d)) cv.d<-NA else {
11       n<-length(datos)
12       datos.i<-matrix(datos,n,n);
13       diag(datos.i)<-0 ;
14       lambda.i<-sapply(1:n, function(i) {lambda.t(datos[
15         i],datos.i[,i],h)})
16       cv2.d<-sum(lambda.i,na.rm=TRUE)
17       cv.d<-(cv1.d-2*cv2.d)
18     }
19     return(cv.d) # the CV-score
20   }
21   n<-length(datos)
22   h.min<-2*(datos[n]-datos[1])/n

```

```
20   h.max<-(datos[n]-datos[1])
21   nb<-100
22   grid.h<-seq(h.min,h.max,length=nb)
23   grid.cv<-sapply(grid.h,CV)
24   plot(grid.h,grid.cv)
25   hcv<-grid.h[which.min(grid.cv)]
26   return(list(hcv=hcv,cv.score=grid.cv))
27 }
```

3.2.3. Ejemplo práctico

Hemos construido un ejemplo simulado, considerando como modelo subyacente la función ROCOF con forma log-lineal. Para examinar el efecto del tamaño muestral, hemos considerado 3 casos distintos: $n = 100, 500, 1000$. Para cada tamaño muestral hemos simulado una muestra y hemos obtenido el ajuste local constante detallado en la Sección 3.1.1. Además hemos usado el código presentado arriba para obtener en cada caso un valor óptimo para el ancho de banda usando el método de validación cruzada que se explica en la Sección 3.2. Los resultados se presentan en la Figura 3.2.3. De izquierda a derecha presentamos los resultados basados en muestras de tamaño 100, 500 y 1000 respectivamente. Puede apreciarse a partir de los gráficos, como es de esperar, que el ajuste es más preciso conforme aumenta el tamaño de la muestra. Además el ancho de banda estimado óptimo es más pequeño conforme aumentan los datos. Cuanto menos datos tenemos, es preciso aumentar el tamaño de la ventana alrededor del punto de estimación para tener información suficiente. Puede apreciarse cómo el ajuste empeora conforme nos vamos acercando a la cola derecha de la distribución. Se produce lo que se conoce como efecto frontera. Si bien este problema se va aliviando conforme aumenta el número de observaciones, sería conveniente utilizar estimadores tipo núcleo de orden mayor como por ejemplo el estimador local lineal. En puntos de estimación del interior del recorrido de la variable, los estimadores local constante (implementado en este trabajo) y local lineal presentan un comportamiento similar y las mismas propiedades asintóticas. En cambio, cuando se trata de estimar en puntos de la frontera, el estimador basado en aproximaciones lineales localmente presenta un rendimiento superior comparado con el caso que estudiamos aquí.

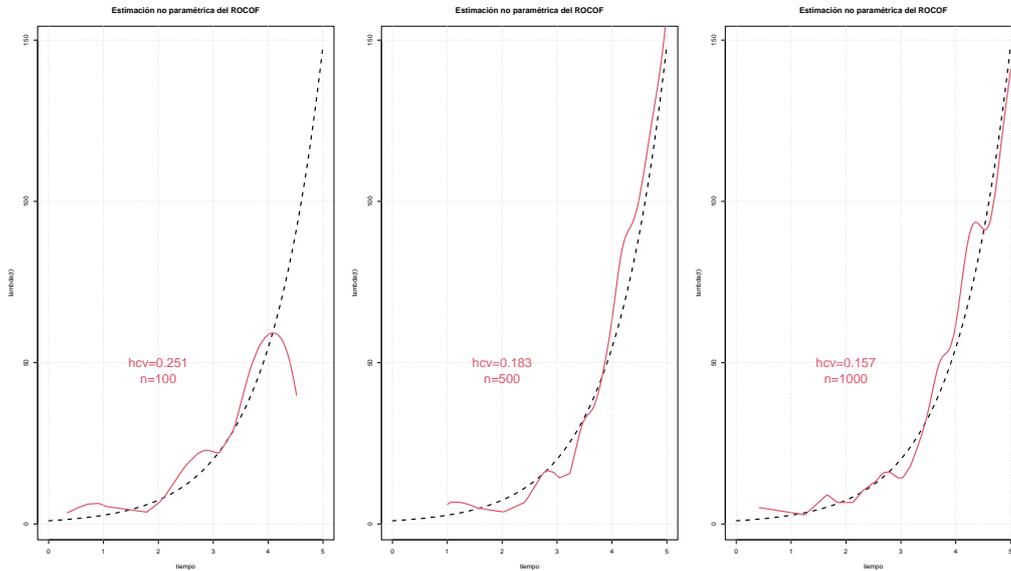


Figura 3.1: Estimación no paramétrica de la función ROCOF con muestras simuladas con distinto tamaño muestral desde el modelo log-lineal. Ancho de banda seleccionado por validación cruzada.

3.3. Bandas de confianza para la función ROCOF

3.3.1. Descripción del modelo

Al ocuparse del problema de la estimación de una cantidad desconocida, se debe tener en cuenta cómo cuantificar la incertidumbre en la estimación. Si el objetivo es estimar un único valor numérico, entonces, se construye un intervalo de confianza para dar una idea de la precisión de una estimación puntual. En el contexto de la estimación de curvas, las bandas de confianza deben desempeñar una función similar.

En esta sección se consideran las técnicas ideadas por Cowling *et al.* (1996) y (1996) y utilizadas en Phillips (2000, 2001) para conseguirlo. En el artículo seminal de Cowling *et al.* (1996), se presentan varios métodos bootstrap para construir regiones de confianza para la función de intensidad de un PPNH. Además, estos métodos son aplicados por Phillips (2000, 2001) a la estimación del ROCOF esperado de un sistema reparable.

Se define ahora lo que se entiende por bandas de confianza.

Definición 3.1. Región de confianza Un conjunto aleatorio conectado no vacío \mathcal{B} del rectángulo $[0, \tau] \times [0, \infty)$ tal que $\mathcal{B} \cap [0, \tau] \times [0, \infty)$ es no vacía para cada $t \in [0, \tau]$ se denomina región de confianza para λ sobre el conjunto $\mathcal{I} \subseteq [0, \tau]$ con cobertura $1 - \alpha$ si

$$Pr \{(t, \lambda(t)) \in \mathcal{B} \text{ para todo } t \in \mathcal{I}\} = 1 - \alpha$$

En el caso de que \mathcal{I} sea un punto o un conjunto finito de puntos, los intervalos de confianza se obtienen utilizando la definición anterior. Cuando se considera que \mathcal{I} es un intervalo, si derivamos lo que se conoce como banda de confianza simultánea, que puede entenderse como una colección de intervalos de confianza para todos los valores de $t \in \mathcal{I}$, construidos para tener probabilidad de cobertura simultánea de, digamos, $1 - \alpha$. En otras palabras, todos los intervalos de la colección cubren su valor verdadero simultáneamente con una probabilidad de $1 - \alpha$.

Consideramos que \mathcal{I} es un intervalo, por lo que construiremos una banda de confianza simultánea que, para simplificar, se representará como $\hat{\lambda}(t) \pm B(t)$, y que, según la definición antes dada, debe satisfacer la condición

$$Pr \{\hat{\lambda}(t) - B(t) \leq \lambda(t) \leq \hat{\lambda}(t) + B(t), \forall t \in \mathcal{I}\} = 1 - \alpha$$

A continuación, se considera las bandas de confianza del bootstrap percentil-t, tal como se presenta en Cowling *et al.* (1996). Ahora, se dan los pasos para derivar las bandas de confianza para el ROCOF siguiendo las pautas sugeridas por estos autores. En primer lugar, se presentan uno de los métodos de remuestreo introducidos en el trabajo citado, donde se aplican los métodos bootstrap en el contexto del PPNH. En consecuencia, consideramos un procedimiento percentil-t para construir regiones de confianza.

Sean $0 < S_1 < S_2 < \dots < S_n < \tau$ los tiempos de fallo observados en el intervalo $[0, \tau]$.

3.3.2. Método de remuestreo

En función de los datos, generar n^* a partir de una distribución de Poisson con parámetro $\Lambda(\tau) = \int_0^\tau \lambda(u) du$ y extraer $S_1^*, S_2^*, \dots, S_{n^*}^*$ mediante un muestreo aleatorio con reemplazo n^* veces del conjunto de datos S_1, S_2, \dots, S_n . Este método produce remuestreos de tamaños variables y que pueden incluir empates. El procedimiento está motivado por el hecho de que, condicionado al evento $N(\tau) = n$, las muestras $S_1 < S_2 < \dots < S_n$ son los estadísticos ordenados de una muestra de tamaño n extraída de una distribución con pdf $f(t) = \lambda(t)/\Lambda(\tau)$. La versión bootstrap del estimador $\hat{\lambda}$ es entonces

$$\hat{\lambda}_h^*(t) = \frac{1}{h} \sum_{i=1}^{n^*} K\left(\frac{S_i^* - t}{h}\right)$$

Una razón para elegir este procedimiento de remuestreo en lugar de otros es que, en este caso, se cumple que, $E_*[\hat{\lambda}_h^*] = \hat{\lambda}_h$, lo que simplifica algunas expresiones, como se comprobará más adelante. Aquí, el subíndice * indica que la expectativa se calcula condicionada al conjunto de datos.

3.3.3. Aproximación Bootstrap

El método que se considera para construir bandas de confianza simultáneas para λ se basa en la aproximación bootstrap

$$\begin{aligned} X^*(t) &= \frac{\hat{\lambda}_h^*(t) - E_*[\hat{\lambda}_h^*(t)]}{\sqrt{\text{Var}(\hat{\lambda}_h(t))}} \\ &= \frac{\hat{\lambda}_h^*(t) - \hat{\lambda}_h(t)}{\sqrt{\text{Var}(\hat{\lambda}_h(t))}}, t \in [0, \tau] \end{aligned}$$

de

$$X(t) = \frac{\hat{\lambda}_h(t) - E[\hat{\lambda}_h(t)]}{\sqrt{\text{Var}(\hat{\lambda}_h(t))}}, t \in [0, \tau]$$

En Cowling *et al.* (1996) se utiliza la siguiente expresión como aproximación bootstrap

$$X^*(t) = \frac{\hat{\lambda}_h^*(t) - \hat{\lambda}_h(t)}{\sqrt{\hat{\lambda}_h^*(t)}}, t \in [0, \tau]$$

3.3.4. Bandas de confianza

Se considera aquí un procedimiento para construir bandas de confianza para λ de una amplitud proporcional a la desviación estándar. En concreto, se define C_α de forma que

$$Pr\{|X^*(t)| \leq C_\alpha, \forall t \in \mathcal{I}\} = 1 - \alpha$$

donde, como en el caso anterior, el subíndice * indica que la probabilidad se calcula condicionada al conjunto de datos. Así, para cualquier $t \in [0, \tau]$, considere los límites dados por

$$\mathcal{B} = \max \left\{ 0, \hat{\lambda}_h(t) - C_\alpha \sqrt{\hat{\lambda}_h(t)} \right\}$$

y

$$\mathcal{B} = \hat{\lambda}_h(t) + C_\alpha \sqrt{\hat{\lambda}_h(t)}$$

y definir la región

$$\mathcal{B} = [0, \tau] \times (\mathcal{B}_1, \mathcal{B}_2)$$

Como se puede apreciar, en cada $t \in [0, \tau]$, la amplitud del intervalo correspondiente es proporcional a $\sqrt{\hat{\lambda}_h(t)}$, lo que refleja la variabilidad de la estimación puntual.

3.3.5. Implementación en R

Incluimos en esta sección un código original implementado en R para la construcción de bandas de confianza para la función ROCOF tal como se explica en secciones anteriores.

```

1  Bandas de confianza usando re-muestreo
2  conf.bands<-function(datos,B=200,alpha=0.05,h)
3    #B es el numero de muestras bootstrap
4  {
5    h<-h.cv(datos)$hcv
6    t.max<-max(datos)
7    t.grid<-datos;#seq(0,t.max,length=101)
8    n.grid<-length(t.grid)
9    #estimador basado en la muestra original
10   lambda.0<-lambda.t(t=t.grid,datos=datos,h=h)
11   Lambda.max<-Lambda.t(t=t.max,datos=datos,h=h)
12   #Generamos un valor desde Poisson(Lambda.max)
13   n.boot<-rpois(B,Lambda.max)
14   #obtenemos una muestra bootstrap con reemplazamiento
15   # desde la muestra original y de tama\~no n.boot
16   #datos.boot<-sort(sample(datos,n.boot,replace=TRUE))
17   #creamos una lista con todas las muestras bootstrap
18   datos.boot<-sapply(n.boot,sample,x=datos,replace=TRUE)
19   #ordenamos las muestras
20   datos.boot<-lapply(datos.boot,sort)
21   h.boot<-double(B) #vector que contiene el h.cv de cada
      muestra bootstrap

```

```

22  #h.boot<-lapply(datos.boot,h.cv) ###
23  h.boot<-rep(h,B)  ## no estamos interesados en el ancho
      de banda en este problema
24  #Calculamos las versiones bootstrap del estimador mu
25  lambda.boot<-matrix(0,n.grid,B)
26  x.boot<-matrix(0,n.grid,B)
27  for (b in 1:B)
28  {
29      lambda.boot[,b]<-lambda.t(t=t.grid,datos=datos.boot
      [[b]],h=h.boot[[b]])
30      x.boot[,b]<-abs((lambda.boot[,b]-lambda.0)/sqrt(
      lambda.boot[,b]))
31  }
32  #Percentile-t confidence bands
33  #ordenamos las aproximaciones x.boot en funcion del
      valor minimo
34  mini.boot<-sapply(x.boot,min)
35  n.alpha<-round(B*(alpha/2))
36  C.alpha<-mini.boot[n.alpha]
37  #Construimos las bandas de confianza
38  B.inf<-lambda.0-C.alpha*sqrt(lambda.0)
39  ceros<-which(B.inf<0)
40  B.inf[ceros]<-0
41  B.sup<-lambda.0+C.alpha*sqrt(lambda.0)
42  #RESULTADO
43  bands<-data.frame('t.grid'=t.grid,'B.inf'=B.inf,'
      lambda.0'=lambda.0,'B.sup'=B.sup)
44  bands
45  }

```

Por último, hemos considerado también el modelo log-lineal para ilustrar la construcción de bandas de confianza alrededor de la función ROCOF que hemos implementado y hemos presentado en esta sección. En este caso, hemos considerado una muestra de tamaño intermedio $n = 500$ y hemos construido intervalos de confianza al nivel 0.95. Hemos considerado un total de 500 muestras bootstrap para cálculo de la precisión de los intervalos siguiendo el código anterior. Como puede verse las bandas de confianza heredan de alguna manera el problema de falta de ajuste en la frontera de la distribución. Los resultados se presentan en la Figura 3.3.5.

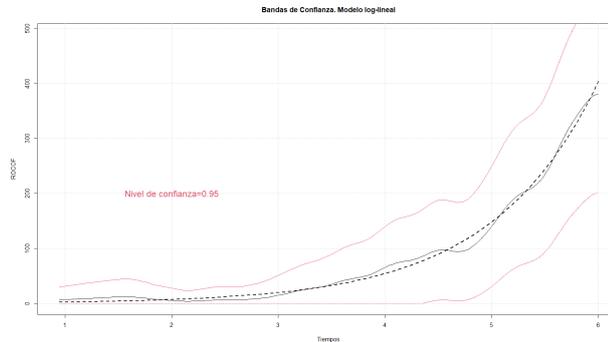


Figura 3.2: Bandas de confianza al nivel 0.95 para el modelo log-lineal a partir de una muestra de tamaño 500 y con ancho de banda seleccionado por validación cruzada.

3.4. Estimación con múltiples realizaciones

3.4.1. Descripción del modelo

Se tratará la estimación de la función de intensidad de un PPNH en el caso de que se disponga de varias realizaciones del proceso. En el lenguaje de la configuración de la fiabilidad, suponemos que dos o más sistemas idénticos están siendo observados independientemente a través del tiempo. Supongamos que se tienen datos de fallos de una colección de k sistemas idénticos. La aparición de fallos en cada uno de ellos puede modelarse como un proceso de Poisson no homogéneo con una función de intensidad desconocida $\lambda(t)$. Así, la observación de cada sistema individual proporciona una realización de un PPNH con intensidad $\lambda(t)$. Sea $N_j(t)$ que denota el número acumulado de fallos del sistema j antes del tiempo t . Seguimos a cada sistema de la muestra durante un periodo de tiempo que se supone independiente de la historia del sistema. Durante el intervalo de tiempo $[0, \tau_j]$, el sistema j th presenta fallos en los momentos $S_{j1} < S_{j2} < \dots < S_{jn_j} \leq \tau_j$, para cada $j = 1, 2, \dots, k$.

En Bartoszyński *et al.* (1981) desarrollan un estudio del mecanismo cinético para la propagación del cáncer. Modelan la aparición de metástasis como un proceso de Poisson no homogéneo con intensidad λ cuya forma funcional no se especifica de antemano, sino que los autores dejan que los datos determinen su forma. Así, consideran procedimientos de estimación no paramétricos para aproximar el valor de λ , en particular presentan una estimación de tipo ventana para la función de intensidad. Adaptamos su razonamiento para estimar la ROCOF de un proceso de recurrencia de fallos en el contexto

de los sistemas de fiabilidad. Se define el siguiente estimador

$$\hat{\lambda}_h(t) = \frac{1}{h} \frac{\#\text{fallos en } (t - h/2, t + h/2)}{\#\text{sistemas en observacion en el momento } t}$$

Sea $\delta_j(t) = I[\tau_j \geq t]$ el indicador de riesgo correspondiente al sistema j th de la muestra. En otras palabras, $\delta_j(t) = 1$ si el sistema j sigue siendo observado en el momento t y $\delta_j(t) = 0$ en caso contrario. Sea $\delta.(t) = \sum_{j=1}^k \delta_j(t)$ el tamaño del conjunto de riesgo en el momento t . El estimador $\hat{\lambda}_h(t)$ definido antes, puede escribirse como

$$\hat{\lambda}_h(t) = \frac{1}{h} \sum_{j=1}^n \frac{\delta_j(t)}{\delta.(t)} \sum_{i=1}^{n_j} I[S_{ji} \in (t - h/2, t + h/2)].$$

Se trata de una estimación de tipo núcleo con función núcleo $K(\cdot) = I[|u| \leq 0,5]$.

Entonces, definiendo $K_h(\cdot) = K(\cdot/h)/h$, el estimador $\hat{\lambda}_h(t)$ se puede escribir como

$$\hat{\lambda}_h(t) = \sum_{j=1}^n \frac{\delta_j(t)}{\delta.(t)} \sum_{i=1}^{n_j} K_h(S_{ji} - t)$$

y otras funciones núcleo diferentes de $K(\cdot)$, podrían usarse. Sin embargo, dado que en el entorno de la estimación del núcleo se acepta generalmente que la elección de la forma del núcleo tiene una importancia relativa (en comparación con la elección del parámetro de ancho de banda h), nos limitaremos a la función dada $K(\cdot)$. Obsérvese que la expresión $\hat{\lambda}_h(t) = \sum_{j=1}^n \frac{\delta_j(t)}{\delta.(t)} \sum_{i=1}^{n_j} K_h(S_{ji} - t)$ se obtiene como la media de las estimaciones del kernel obtenidas al considerar cada realización del PPNH por separado.

Como se establece en Bartoszyński *et al.* (1981), si el número de sistemas censurados en el intervalo $(t - h/2, t + h/2)$ es despreciable en relación con $\delta.(t)$ y si λ es dos veces diferenciable en t , entonces, dado el tamaño del conjunto de riesgo, se puede derivar la siguiente expresión para el error cuadrático medio (*MSE*)

$$\begin{aligned} MSE_{\delta}(t, h) &= E_{\delta} \left[\left(\hat{\lambda}_h(t) - \lambda(t) \right)^2 \right] \\ &= \frac{\lambda(t)}{\delta} + \frac{h^2 \lambda^4(t)}{4} + O(\delta.(t)^{-1}) + O(h^3). \end{aligned}$$

Esta expresión se puede minimizar, dando lugar al siguiente valor del ancho de banda que se considerará óptimo en términos de *MSE*

$$h_{opt}(t) = \left(\frac{2}{\delta \cdot(t)} \right)^{1/3} \frac{1}{\lambda(t)}.$$

En Gámiz *et al.* (2011) adaptan el algoritmo iterativo presentado en Bartoszynski *et al.* (1981) para estimar $\lambda(t)$ en cada momento de fallo. El algoritmo puede resumirse en unos pocos pasos.

Algoritmo. Estimación de ROCOF con tiempos de fallo recurrentes

Paso 1. Considerar como valor inicial de la ROCOF una estimación constante dada por el número total de fallos dividido por la duración total del periodo de observación, es decir

$$\hat{\lambda}_0(t) = \frac{\sum_{j=1}^k n_j}{\max\{\tau_j : j = 1, 2, \dots, k\}}$$

Paso 2. Dada una estimación previa $\hat{\lambda}_i$, calcula h_{i+1} mediante la ecuación $h_{opt}(t) = \left(\frac{2}{\delta \cdot(t)} \right)^{1/3} \frac{1}{\lambda(t)}$ como

$$h_{i+1} = \left(\frac{2}{\delta \cdot(t)} \right)^{1/3} \frac{1}{\hat{\lambda}_i(t)}.$$

Paso 3. Actualizar la estimación de la ROCOF mediante $\hat{\lambda}_{i+1}$ calculado como con $\hat{\lambda}_h(t) = \frac{1}{h} \sum_{j=1}^n \frac{\delta_j(t)}{\delta \cdot(t)} \sum_{i=1}^{n_j} I[S_{ji} \in (t - h/2, t + h/2)]$ con h_{i+1} , y detenerse si este valor actualizado se acerca lo suficiente (según un criterio de convergencia adecuado) al obtenido en la iteración anterior, de lo contrario se pasa a la etapa 2.

3.4.2. Implementación en R

En esta sección incluimos la implementación en R del algoritmo explicado en la sección anterior.

```

1  ## Los datos estan en forma de lista.
2  ## Cada elemento de la lista es una realizacion del
   proceso
3  ## calculamos el estimador para cada realizacion por
   separado
4
5  lambda.t <- function(datos, h, t0) {
```

```

6     K<-function(x) return(dnorm(x,0,1))
7     m<-length(t0)
8     lambda<-sapply(1:m,function(i) {(1/h)*sum(K((t0[i]-
9         datos)/h))})
10    return(lambda)
11  }
12  delta<-function(datos,t0) #k es el numero de
13    trayectorias
14    #definimos la funcion indicadora de censura
15    #para cada trayectoria
16  {
17    indica<-function(v,t) {ifelse( max(v) >= t, 1, 0 )}
18    k<-length(datos)
19    Delta<-double(k) #calculamos el vector de deltas en un
20      t0 particular
21    for (j in 1:k)
22    {
23      Delta[j]<-indica(datos[[j]],t0)
24    }
25    delta.<-sum(Delta)
26    res<-list('delta.j'=Delta, 'delta.'=delta.)
27    res
28  }
29  #####ALGORITMO#####
30  ## Funciones auxiliares.
31  #1. Calculamos el valor inicial del estimador de lambda,
32  # es una estimacion empirica, se trata de una funcion
33  constante
34  lambda.ini<-function(datos)
35  {
36    k<-length(datos)
37    n.j<-double(k) #numero de observaciones por sistema
38    tau.j<-double(k) #tiempo maximo de observacion de cada
39      sistema
40    for (j in 1:k)
41    {
42      n.j[j]<-length(datos[[j]])
43      tau.j[j]<-datos[[j]][n.j[j]]
44    }
45    lambda0<-(sum(n.j))/(max(tau.j))

```

```

44     lambda0
45 }
46 #####
47 #2. Obtenemos el estimador combinado del ROCOF segun
      promedio de las
48 #   estimaciones en las trayectorias individuales
49
50 lambda.avg<-function(datos,h0,t0)
51 {
52     delta.t0<-delta(datos,t0)
53     delta.j<-delta.t0$delta.j
54     Delta<-delta.t0$delta.
55
56     #definimos un vector con la estimacion de lambda para
      cada trayectoria
57     #evaluando la funcion lambda.t:
58
59     k<-length(datos)
60     v.lambda<-double(k)
61
62     for(j in 1:k) {v.lambda[j]<-lambda.t(datos[[j]],h0,t0)
      }
63
64     lambda.av<-sum(delta.j*v.lambda)/Delta
65     lambda.av
66 }
67 #####
68
69 #3. Aproximamos el valor del error cuadratico medio
      segun expresion en
70 #   Bartozynski para usarlo mo criterio de parada
71
72 mse<-function(lambda,delta,h)
73 {
74     mse<-lambda/(h*delta)+(h^2*lambda^4)/4
75     mse
76 }
77
78 #4. Pasos intermedios del algoritmo
79 recloc<-function(datos,t0)
80 {
81     #actualizamos el conjunto de riesgo en t0
82     delta.t0<-delta(datos,t0)

```

```

83  delta.j<-delta.t0$delta.j
84  Delta<-delta.t0$delta.
85
86  #valores iniciales
87  lambda.0<-lambda.ini(datos)
88  ce.0<-lambda.0
89  h.0<-((2/Delta)^(1/3))*(1/lambda.0)
90  mse.0<-mse(lambda.0,Delta,h.0)
91
92  # primera iteracion
93  # iter<-1
94  # lambda.1<-lambda.avg(datos,h.0,t0)
95  # h.1<-((2/Delta)^(1/3))*(1/lambda.1)
96  # mse.1<-mse(lambda.1,Delta,h.1)
97  #
98  #construimos vectores que guardan los calculos hasta
99  # el momento
100 v.h<-c(h.0)
101 v.lambda<-c(lambda.0)
102 v.mse<-c(mse.0)
103 v.iter<-c()
104 v.cesar<-c(ce.0)
105 differ<-1;v.dif<-c()
106 iter<-0
107 while (differ>0.001)
108 {
109   #actualizamos los calculos
110   iter<-iter+1
111   lambda.nuevo<-lambda.avg(datos,v.h[iter],t0)
112   h.nuevo<-((2/Delta)^(1/3))*(1/lambda.nuevo)
113   mse.nuevo<-mse(lambda.nuevo,Delta,h.nuevo)
114
115   #actualizamos los vectores
116   v.h<-c(v.h,h.nuevo)
117   v.lambda<-c(v.lambda,lambda.nuevo)
118   v.mse<-c(v.mse,mse.nuevo)
119   v.cesar<-c(v.cesar,mean(v.lambda))
120   differ<-abs(v.cesar[iter+1]-v.cesar[iter])
121   v.dif<-c(v.dif,differ)
122
123   #print(cbind(v.h,v.lambda,v.mse))
124 }
resultados<-data.frame(v.h,v.lambda,v.mse)

```

```

125     colnames(resultados)<-c("h","lambda","mse")
126     resultados
127 }
128 ## Funcion que ejecuta el algoritmo:
129 # Calculamos el valor final del estimador en una malla
130 # en [0, tau]
131 # Las observaciones son los tiempos de ocurrencias
132 lambda.mult<-function(datos)
133 {
134     red<-seq(range(datos)[1],range(datos)[2],length=100)
135     m<-length(red)
136
137     lambda.red<-lapply(red,recloc,datos=datos)
138     lambda.loc<-double(m)
139     for (i in 1:m)
140     {
141         n<-nrow(lambda.red[[i]])
142         lambda.loc[i]<-lambda.red[[i]]$lambda[n]
143     }
144     ## presentamos el resultado en un grafico:
145     plot(red,lambda.loc,main=paste('Datos recurrentes.
146     Muestra simulada. Modelo:',modelo,sep=''),xlab='
147     Time',ylab='ROCOF')
148 }

```

3.4.3. Ejemplo práctico

Se simulan diez trayectorias muestrales de tamaño entre 50 y 70 obtenidos de forma aleatoria. Se obtuvo que las diez trayectorias tienen de forma respectiva, tamaños 62, 69, ..., 67 como se muestra. A continuación se muestran los sucesivos tiempos de fallo de las diez máquinas, teniendo que, para la máquina 1 los tiempos de fallo son 0,1381353, 0,2125607, ..., 34,1369655 y de forma sucesiva para las otras nueve máquinas. El modelo que se utiliza es el periódico, la función *coseno*. Se compara el gráfico de la función ROCOF (función oscilante) con el gráfico de las trayectorias generadas; aunque el gráfico es oscilante y se parece al gráfico del ROCOF, hay puntos muy dispersos. Los resultados gráficos del ajuste se presentan en la Figura 3.4.3.

```

> nn
62 69 66 66 59 69 55 68 70 67

```

List of 10

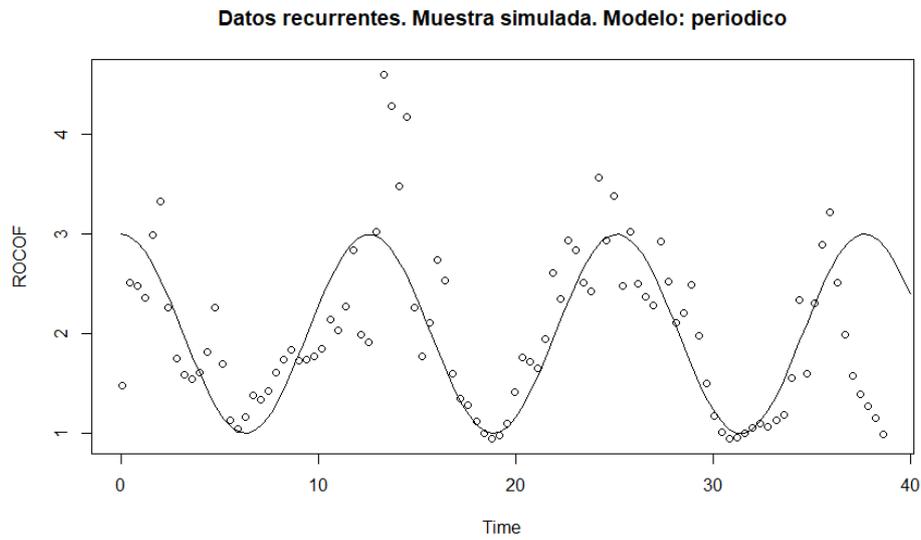


Figura 3.3: Ajuste con múltiples trayectorias muestrales

```

$ : num [1:62] 0.138 0.213 0.75 0.75 1.249 ...
$ : num [1:69] 0.0489 0.5209 0.6414 0.9003 0.943 ...
$ : num [1:66] 0.513 0.727 0.792 2.031 2.369 ...
$ : num [1:66] 0.687 0.981 1.66 1.78 2.014 ...
$ : num [1:59] 0.193 0.215 0.384 0.54 1.315 ...
$ : num [1:69] 0.24 0.515 0.978 1.317 1.41 ...
$ : num [1:55] 0.0746 0.1768 0.3115 0.4683 1.6298 ...
$ : num [1:68] 0.792 1.308 1.964 1.995 2.267 ...
$ : num [1:70] 1.5 1.91 2.07 2.32 3.6 ...
$ : num [1:67] 0.136 0.99 1.321 1.362 1.625 ...
    
```

Capítulo 4

Aplicación Práctica

4.1. Descripción general de los datos

A lo largo de los años se ha visto cómo las maquinarias o componentes de esas maquinarias sufren discontinuidades o defectos; estos pueden aparecer a través del proceso de fabricación y/o según las condiciones de funcionamiento. Con el tiempo, los defectos estimulan la aparición de grietas, haciendo que éstas originen gran concentración de esfuerzos que hacen que haya fallas en las maquinarias. Para la comunidad científica, se vuelve interesante investigar acerca del efecto del esfuerzo aplicado, el tamaño de las grietas en las piezas, la forma de la grieta y, la velocidad con que crece esa grieta. En virtud de la importancia de este tema, se hará un análisis del comportamiento de estos datos, si son recurrentes, si llevan un proceso de Poisson, si se les puede hacer algún ajuste con modelos específicos, etcétera.

Se toma un conjunto de datos que fueron recopilados por el profesor BM Hillberry, y que fueron publicados en Virkler et al. (1979), con el tema *La naturaleza estadística de la propagación de grietas por fatiga*, así mismo, tomando como fuente a Eric J. Tuegel. En su momento se realizó 68 pruebas repetidas con una amplitud constante en una aleación de aluminio para investigar la propagación de grietas por fatiga. En cada uno de estos ensayos, se puede observar el número de ciclos que conduce a longitudes de fisura fijas. Contra el supuesto natural de que algo se observa en momentos fijos, en Virkler el tiempo es la variable dependiente y la longitud de la grieta (milímetro) es la variable independiente. Por tanto, la longitud de la grieta se trató aquí como un vector de tiempo t . Sin embargo, para nuestro estudio, se tomará al vector de tiempo t , como la variable independiente, cuya variable *tiempo* se mide en ciclos de funcionamiento, partiendo de 0 ciclos hasta llegar a menos

de 30,000 ciclos; mientras que, la longitud de la grieta será la variable dependiente. Modelizamos el fenómeno usando un proceso de Poisson $\{N(t)\}$, donde $N(t)$ tiene un salto de tamaño 1 y cada vez que la grieta se incrementa una cantidad fija de $0,2mm$, siendo este el suceso de interés, entendemos que la grieta aumenta de modo continuo, pero, definimos el proceso como si el sistema recibiese golpes en momentos determinados que incrementan la grieta en esa magnitud ($0.2mm$ cada golpe). La matriz muestra 69 columnas, donde la primera representa el tamaño de la grieta en milímetros, podemos observar en cada fila incrementos de $0,2mm$, y, la segunda columna es una muestra que representa la serie de tiempos observados en ciclos, es decir, el número de ciclos en que se detecta el incremento de la longitud de esa grieta, teniendo en total 163 fallos y S_n será el número de ciclos hasta cada tiempo t . En cada columna se tendrá una muestra, en total tenemos 68 muestras. En la primera muestra, en el tiempo 1 el valor de S_1 es de 5529 ciclos, en el tiempo 2, $S_2 = 10408$, en el tiempo 3, $S_3 = 15408, \dots$. El primer fallo ocurrió después de 5529 ciclos, de manera sucesiva hasta llegar al último fallo registrado, después de 237293 ciclos, con una grieta de $49,8mm$. En la muestra 2 (columna 3) el primer fallo ocurre después de 6232 ciclos, y el último fallo registrado fue después de 242586. Con esto, podemos decir que, tenemos 68 sistemas a los cuales les llegan golpes según un proceso de Poisson. Veamos la siguiente tabla, que contiene los primeros 10 tiempos de fallo de los 163 en el primer sistema.

Número de fallo i	Longitud de la grieta (mm)	Tiempo de fallo (S_i)
0	9.0	0
1	9.2	5529
2	9.4	10408
3	9.6	15408
4	9.8	20095
5	10.0	24532
6	10.2	28443
7	10.4	32809
8	10.6	37677
9	10.8	41278
10	11.0	45554

Tabla 4.1: Primeros tiempos de fallo para el sistema 1.

4.2. Ajuste paramétrico

Veremos las trayectorias de todos los sistemas (desde el sistema 1 hasta el sistema número 68). Se tienen datos recurrentes según un proceso de Poisson. Para graficar estas trayectorias, primero, se normalizaron los tiempos de fallo, dividiéndolos por 10000; por tanto, el eje horizontal, t , eje de las abscisas, llega a un máximo de 30 ciclos, mientras que el eje vertical, eje de las ordenadas, $N(t)$, se mantiene hasta máximo 164 número de fallos. En la figura 4.1 se observa que estas curvas representan el número de fallos (incrementos de magnitud 0.2mm) acumulados hasta el tiempo t .

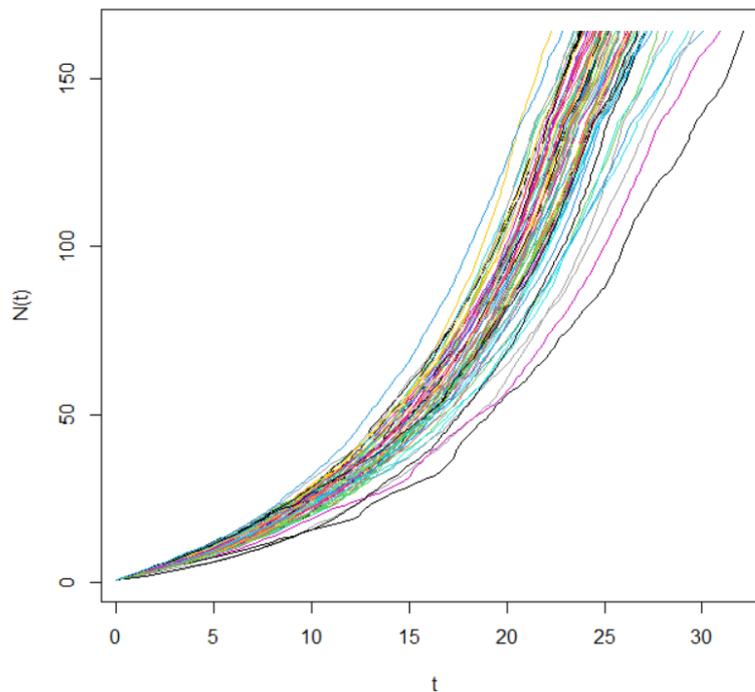


Figura 4.1: Trayectorias muestrales de los 68 sistemas.

En este enfoque paramétrico llevamos a cabo un análisis individual. Es decir presentamos un estudio de la tendencia en una trayectoria (en particular hemos elegido 5 y 28, pero no hay ninguna razón particular para seleccionar estas dos, podríamos hacer el estudio con otras muestras diferentes). Sus gráficos parecen tener tendencia, y llevar una forma exponencial.

Se calcula el test de tendencia de Laplace, para los sistemas 5 y 28 respectivamente. En ambos sistemas se obtuvo un p-valor bastante pequeño, menor

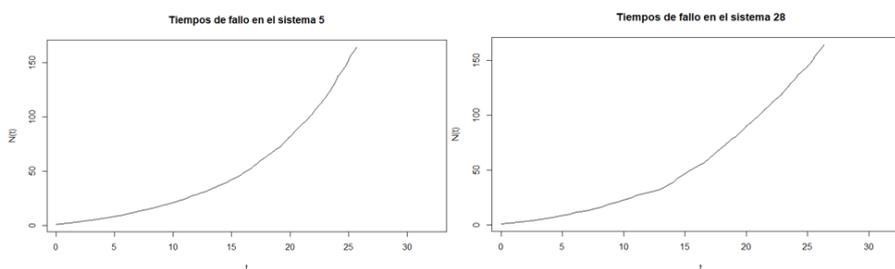


Figura 4.2: Trayectorias de los sistemas 5 y 28.

que 0,005. En el sistema 5 el p-valor es de $7,193424 \times 10^{-22}$ indicando con esto que, se rechaza la tendencia constante. De igual forma sucede con el sistema 28, se rechaza la tendencia constante. Al realizarse el test de Lewis-Robinson, los resultados se mantienen; mientras que, con Military Handbook se acepta la tendencia constante. Recordemos que, en el capítulo 2.6, se menciona que la prueba de Laplace es mejor para encontrar significancia si los datos provienen de un sistema que sigue la ley exponencial y tiene una base similar para comprobar tendencia en el modelo loglineal; mientras que la prueba Military Handbook es mejor para encontrar significancia para un modelo de ley de potencia, donde podemos encontrar diferencias entre los resultados obtenidos al usar ambas pruebas. A continuación se muestran los resultados

Test de tendencia	Test.value	p-valor
Laplace	9.610876	$7.193424e-22$
Military Handbook	140.2076	1
Lewis-Robinson	28.90631	$9.945483e-184$
Test de tendencia	Test.value	p-valor
Laplace	8.334313	$7.79494e-17$
Military Handbook	154.4514	1
Lewis-Robinson	24.15385	$6.802275e-129$

Tabla 4.2: Test de tendencia para los sistemas 5 y 28, respectivamente.

Se hace un ajuste paramétrico con los modelos potencial y exponencial a los datos recopilados del sistema 5 y del sistema 28. Para el sistema 5 se muestran las curvas empíricas y el ajuste obtenido. Observemos la tendencia que llevan las gráficas y aparentemente el ajuste potencial resulta ser más significativo, en comparación con el exponencial; de igual forma con el sistema 28, utilizando los modelos potencial y exponencial respectivamente, las curvas

ajustadas y la curva curva empírica se muestran la figura siguiente. Parece ser que con el modelo potencial, la curva se asemeja mas a la curva empírica.

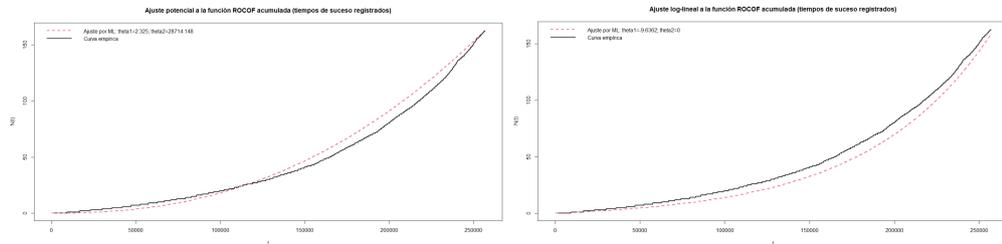


Figura 4.3: Ajuste paramétrico al sistema 5. El panel de la izquierda es el ajuste exponencial y el de la derecha el potencial.

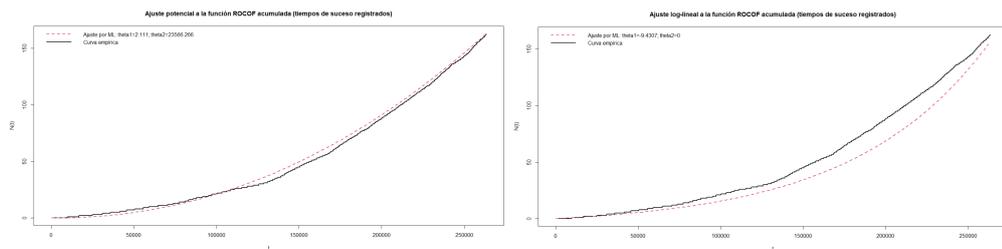


Figura 4.4: Ajuste paramétrico al sistema 28. El panel de la izquierda es el ajuste exponencial y el de la derecha el potencial.

En la figura 4.5 tenemos dos gráficos. En el panel de la izquierda observamos los tiempos entre llegadas del sistema 5: para la primera llegada pasaron 8965 ciclos, entre la llegada uno y dos pasaron 7489 ciclos, entre la llegada dos y tres hay una diferencia de 6427 ciclos, y entre la tercera y cuarta llegada hay 7215 ciclos, etc. Podemos ver en el gráfico que el segundo fallo llegó en un menor número de ciclos que el primer fallo, y que el tercer fallo tarda también un número de ciclos más bajo en comparación con el segundo fallo, mas, para el cuarto fallo no sucede de la misma manera. Se espera que a medida pasa el tiempo(aumentando), los tiempos entre fallos va disminuyendo, ya que la máquina tiene más edad, y en el gráfico sí ocurre este fenómeno (véase la forma decreciente) en la mayoría de casos. Mostramos en la figura 4.5 también en el panel derecho entre los tiempos de fallos y tiempos entre llegadas: se observa que este gráfico es decreciente y hay puntos un poco dispersos.

Cualquier tendencia o patrón observado en estos gráficos alerta sobre la no linealidad del modelo en el sentido de que la función intensidad del proceso de Poisson subyacente no es constante, es decir el proceso no es homogéneo.

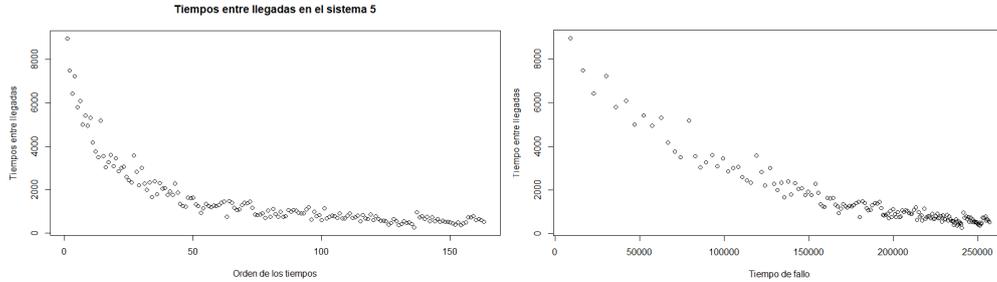


Figura 4.5: Gráfico de tiempo entre llegadas del sistema 5

En la figura 4.6 tenemos al igual que en la figura anterior, dos gráficos que permitirán analizar en este caso al sistema 28. Los tiempos entre llegadas resultaron ser: 9708, 6927, 8276, 6072, ..., de esta manera, vemos que en su mayoría, a medida que la máquina envejece el tiempo entre fallos va disminuyendo, es decir que al aumentar $0,2mm$ de la grieta, para que esto vuelva a ocurrir, pasará un número de ciclos menor que en los anteriores fallos. Aquí también podemos observar la no linealidad del modelo.

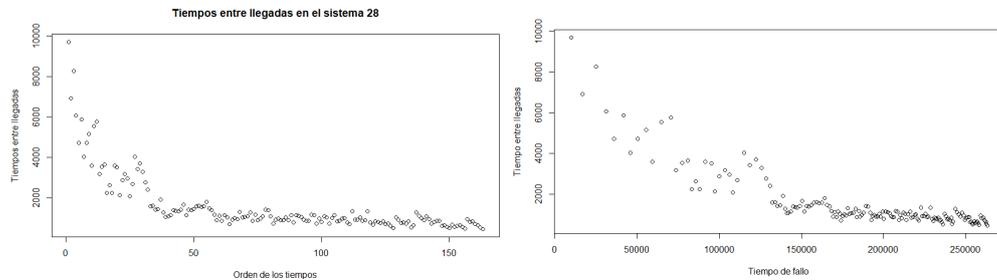


Figura 4.6: Gráfico de tiempo entre llegadas del sistema 5

4.3. Ajuste no paramétrico

4.3.1. Análisis de trayectorias individuales

Se realiza el análisis no paramétrico, lo que significa que se realizará una prueba en la que la distribución de la población no se ajusta a ningún

modelo paramétrico concreto. Se describió antes que las muestras con que se trabajan provienen de la misma distribución y además estas muestras son independientes. Como se introduce en el capítulo 3, se debe observar un sistema o una población de sistemas con las mismas características para lograr la estimación no paramétrica de la ROCOF.

Debido a que la ROCOF es una función suave nos parece adecuado realizar una estimación no paramétrica tipo núcleo de esta función y por lo tanto se debe estimar el parámetro de suavizado, se trabaja con el método de validación cruzada por mínimos cuadrados para encontrar el valor adecuado para el parámetro de ancho de banda, esto explicado en la sección 3.2. Una vez obtenido el parámetro de ancho de banda, se crean las curvas para los 68 sistemas; se reflejan las curvas de los sistemas 5 y 28 (figura 4.7) que representan las funciones score correspondiente al método de validación cruzada tal como se ha explicado en el tema anterior. Como vemos para las dos muestras seleccionadas, se alcanza claramente un mínimo global, lo que determina el valor del ancho de banda óptimo según este criterio que proporcionará la curva que estima la función rocof minimizando el error cuadrático integrado (ISE).

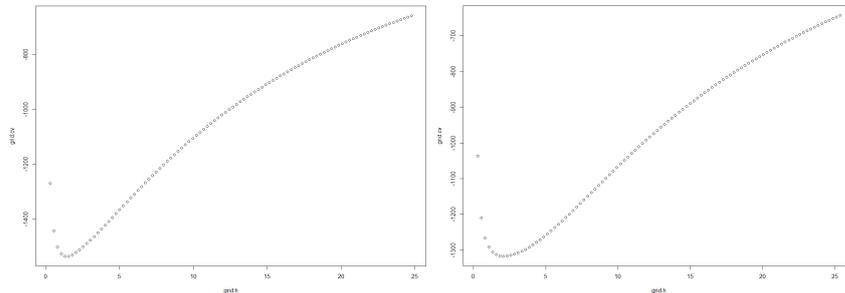


Figura 4.7: Función *score* dada por el criterio de validación cruzada para la selección del ancho de banda para los sistemas 5 y 28.

Ya obtenido el parámetro para el ancho de banda, se calcula y se grafica la estimación no paramétrica de la función ROCOF de los sistemas 5 y 28 (ver figura 4.8), y, en la figura 4.9 se representan las ROCOF para cada muestra y un promedio de todas las curvas. Como posiciones de estimación (en el eje de abscisas), es decir los tiempos en los que hemos estimado, hemos considerado una red de puntos equidistantes entre 0 y el máximo de los tiempos observados a lo largo de las 68 muestras. De manera que todos los ROCOF se han estimado en los mismos puntos de tiempo.

La función ROCOF explica la velocidad con que los fallos se van presentando a lo largo de la recta de tiempos. Vemos que desde el principio esta

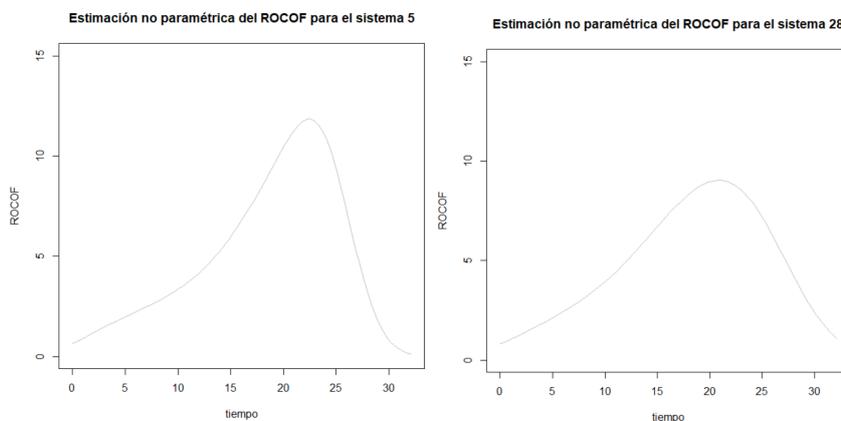


Figura 4.8: Estimación no paramétrica de la función intensidad (ROCOF) para los sistemas 5 y 28.

función crece de modo que los fallos van produciéndose con más asiduidad conforme van acumulándose ciclos de funcionamiento. En los dos sistemas se alcanza un máximo alrededor de los 200000-250000 ciclos, a partir de ese momento la curva decrece rápidamente. Es posible que la tensión que provoca la rotura disminuya cuando se alcanza un determinado umbral, la pieza está muy deteriorada en ese nivel y el mecanismo que produce la grieta aplica menos tensión a la pieza desde ese momento haciendo que los fallos se espacien a partir de ese momento. Podríamos pensar en un modelo paramétrico que sugiera esta curva, considerando que el modelo Log-Normal sería adecuado para valores no tan grandes de t , ya que sería monótona en este caso.

Los anchos de banda de las primeras diez muestras se han obtenido en torno a los valores que se observan en la tabla 4.3:

4.3.2. Estimación de la tendencia con múltiples realizaciones

Se estima la función de intensidad de un PPNH si se dispone de varias realizaciones del proceso. Estamos suponiendo que 68 sistemas idénticos están siendo observados independientemente a través del tiempo y que tenemos datos de fallo de esos 68 sistemas. Entonces, como se indica en el capítulo 3.4 la aparición de fallos en cada uno de ellos puede modelarse como un PPNH con función de intensidad desconocida $\lambda(t)$, y los datos determinan su forma siguiendo el procedimiento de Bartoszyński, considerando procedimientos de estimación no paramétricos para aproximar el valor de la función λ .

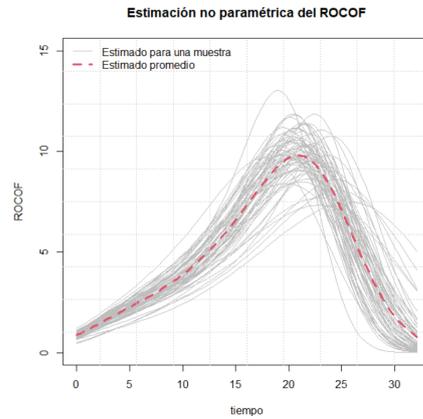


Figura 4.9: Estimación no paramétrica de las tasas de fallo para todos los sistemas considerados individualmente.

Sistema	Ancho de banda
1	3.343537
2	2.938131
3	3.322749
4	3.012232
5	2.58588
6	3.446081
7	2.981913
8	2.856856
9	2.880475
10	4.028929

Se presentan los datos en forma de lista, donde cada elemento de la lista es una realización del proceso y se calcula el estimador constante para cada realización, esto de manera separada. Se calcula el valor inicial del estimador de lambda que es una función constante y se calcula el estimador combinado de la ROCOF según el promedio de las estimaciones en las trayectorias individuales. Luego se calcula el ancho de banda que se considerará óptimo en términos del error cuadrático medio. Por último, se calcula el valor final del estimador en una malla y se presentan los resultados.

Se agrega la curva de la ROCOF integrado (ver figura 4.10). Podemos observar que la función resultante es suave, continua y creciente a lo largo de todo t , es decir que al ir incrementándose el tiempo (ciclos), la frecuencia también va incrementándose, estos sistemas se van deteriorando a medida

pasa el tiempo, algo que nos indicaban los datos, ya que cada vez que había un fallo, había un incremento en el tamaño de la grieta. Un modelo paramétrico que se podría considerar sería el modelo exponencial.

La curva de la figura 4.10, nos da una estimación de cómo se van acumulando los fallos con el paso del tiempo, en este caso de cómo va creciendo la grieta. Si nos fijamos alrededor de los 20000 ciclos de funcionamiento, la curva crece más rápido, esto parece coincidir con el pico de la función no acumulada.

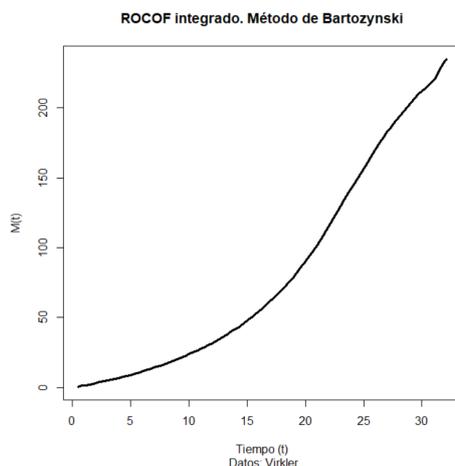


Figura 4.10: ROCOF integrado. Estimación de la tasa de ocurrencia de fallo promedio, combinando la información de todos los sistemas.

Bibliografía

- [1] Andersen, P., Borgan, O., Gill, R. and Keiding, N.: *Statistical Models Based on Counting Processes* (Springer, New York, 1993)
- [2] Ascher, H.: Evaluation of a repairable system reliability using “bad-as-old” concept. *IEEE Transactions on Reliability*, **17**, 103-110 (1968)
- [3] Ascher, H.: Repairable systems reliability, *Encyclopedia of Statistics in Quality and Reliability*, (Ruggeri, F., Kenett, R. and Faltin, F.W. Editors, Wiley, 2007a)
- [4] Ascher, H.: Different insights for improving part and system reliability obtained from exactly same DFOM “failure numbers”. *Reliability Engineering and System Safety*, **92**, 552–559 (2007b)
- [5] Aven, T.: General minimal repair models, *Encyclopedia of Statistics in Quality and Reliability*, (Ruggeri, F., Kenett, R. and Faltin, F.W. Editors, Wiley, 2007)
- [6] Bartoszyński, R., Brown, B.W., McBride, M. and Thompson, J.R.: Some nonparametric techniques for estimating the intensity function of a cancer related nonstationary Poisson process. *The Annals of Statistics*, **9** (5), 1050–1060. (1981).
- [7] Bowman, A. W: *An alternative method of cross-validation for the smoothing of density estimates* . (Biometrika, 1984).
- [8] Brooks, M.M. and Marron, J.S.: Asymptotic optimality of the least-squares cross-validation bandwidth for kernel estimates of intensity functions, *Stochastic Processes and their Applications*, **38**, 157–165 (1991)
- [9] Cinlar, E.: *Introduction to Stochastic Processes*. (Prentice Hall, Englewood Cliffs, NJ, 1975).

- [10] Cowling, A., Hall, P. and Phillips, M.J.: Bootstrap confidence regions for the intensity of a Poisson point process. *Journal of the American Statistical Association*, **91**, 1516–1524 (1996)
- [11] Finkelstein, M.: *Failure Rate Modelling for Reliability and Risk* (Springer, 2008)
- [12] Gámiz, M.L. and Kulasekera, K.B. and Limnios, N. and Lindqvist, B.H.: *Applied Nonparametric Statistics in Reliability* . (Springer, 2011).
- [13] Hollander, M., Samaniego, F.J. and Sethuraman, J.: Imperfect repair, *Encyclopedia of Statistics in Quality and Reliability*, (Ruggeri, F., Kennett, R. and Faltin, F.W. Editors, Wiley, 2007)
- [14] Haussmann, Luciana y Romero, A.C: *Procesos de Poisson no homogéneos aplicados a la extracción de dinero en una red de cajeros automáticos* . (Universidad de la República Uruguay, 2018).
- [15] Krivtsov, V.V.: Practical extensions to NHPP application in repairable system reliability analysis, *Reliability Engineering and System Safety*, **92**, 560–562 (2006)
- [16] Meeker, W. Q. and Escobar, L. A: *Statistical Methods for Reliability Data*. (JOHN WILEY, SONS, INC, 1998).
- [17] Pham, H. and Zhang, X.: NHPP software reliability and cost models with testing coverage. *European Journal of Operational Research*, **145** (2), 443–454 (2003)
- [18] Phillips, M.J.: Bootstrap confidence regions for the expected ROCOF of a repairable system. *IEEE Transactions on Reliability*, **49**, 204–208 (2000)
- [19] Rausand, M. and Høyland A.: *System Reliability Theory. Models, Statistical Methods, and Applications* (John Wiley & Sons, 2004)
- [20] Rudemo, M.: *Empirical choice of histograms and kernel density estimators*. (Scandinavian Journal of Statistics, 1982).
- [21] Rigdon, S.E. and Basu, A.P.: *Statistical Methods for the Reliability of Repairable Systems* (Wiley, 2000)