Fiabilidad

Estadística Industrial

¿Qué es la fiabilidad?

- Permanencia de la calidad de los productos (o servicios) a lo largo del tiempo.
- Capacidad de desarrollar adecuadamente su labor a lo largo del tiempo.

La Calidad .....

- se limita a garantizar que el producto sale de fábrica en buenas condiciones
- ¿Permanece en buenas condiciones?

La Fiabilidad intenta garantizar que el producto permanecerá en buenas condiciones durante un periodo razonable de tiempo.

Calidad vs. Fiabilidad

- Surge la necesidad de considerar un control de calidad basado en el tiempo. El control de calidad habitual, o de inspección, no tiene continuidad temporal: el producto pasa un control o no lo pasa

Herramientas

- **Se estudia mediante el análisis estadístico de datos de supervivencia.**

¿Por qué ESTADÍSTICO?

ISO define fiabilidad como la *probabilidad* de que un componente o sistema, desarrolle durante un periodo de tiempo dado, la tarea que tiene encomendada sin fallos, y en las condiciones establecidas.

Introducción al ADS

- **Vamos a estudiar** *Duraciones de Procesos* **que es algo muy común en muchas ciencias:**
  - **Duración de un componente (Fiabilidad)**
  - **Supervivencia de un paciente a un tratamiento (Medicina)**
  - **Duración del desempleo (Economía)**
  - **Edad de las personas (Demografía y sociología)**

Variables Aleatorias Positivas

En Fiabilidad el tiempo se puede medir de otra manera:

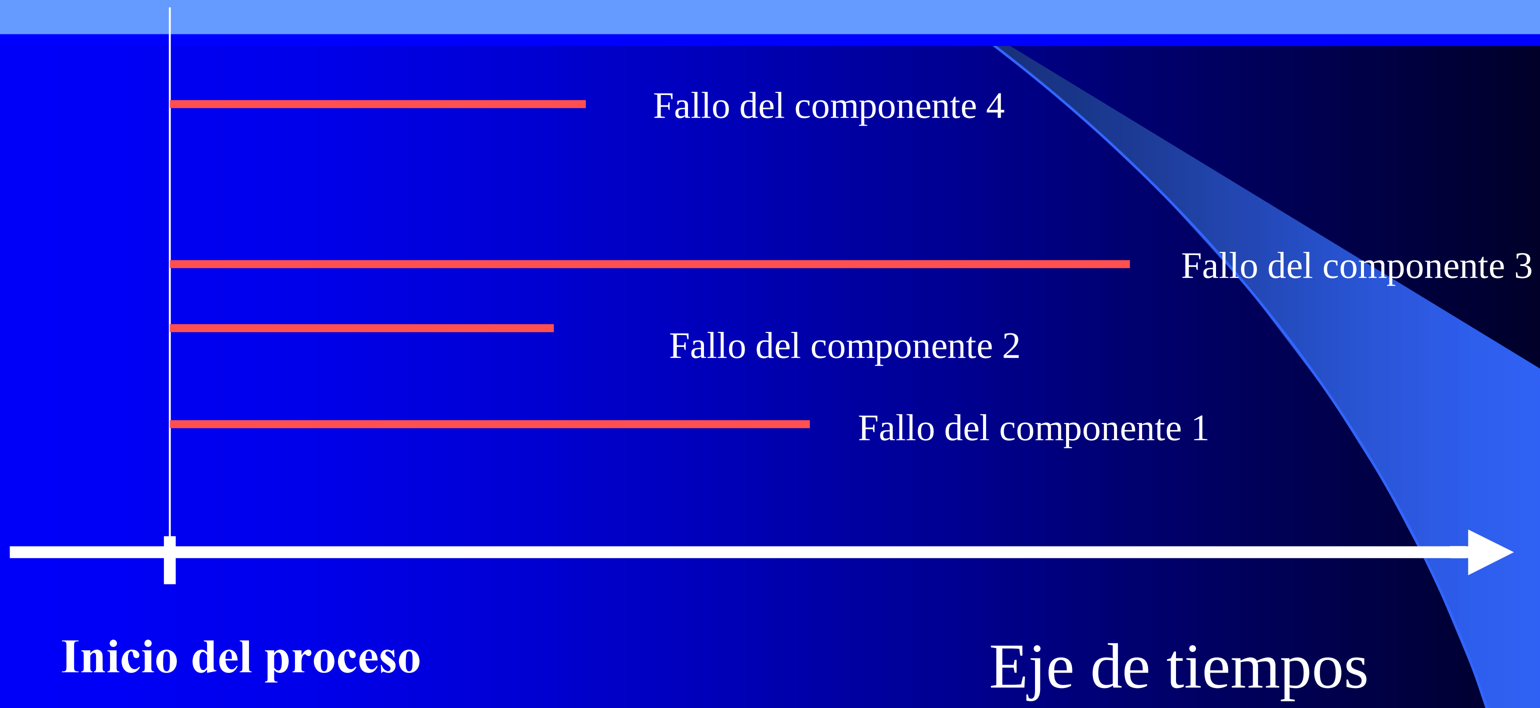
- **Número de veces que se enciende un interruptor.**
- **Ciclos de lavado en una lavadora.**
- **Horas de vuelo de un avión**

Conceptos básicos

Eje de tiempos

Conceptos básicos

Conceptos básicos



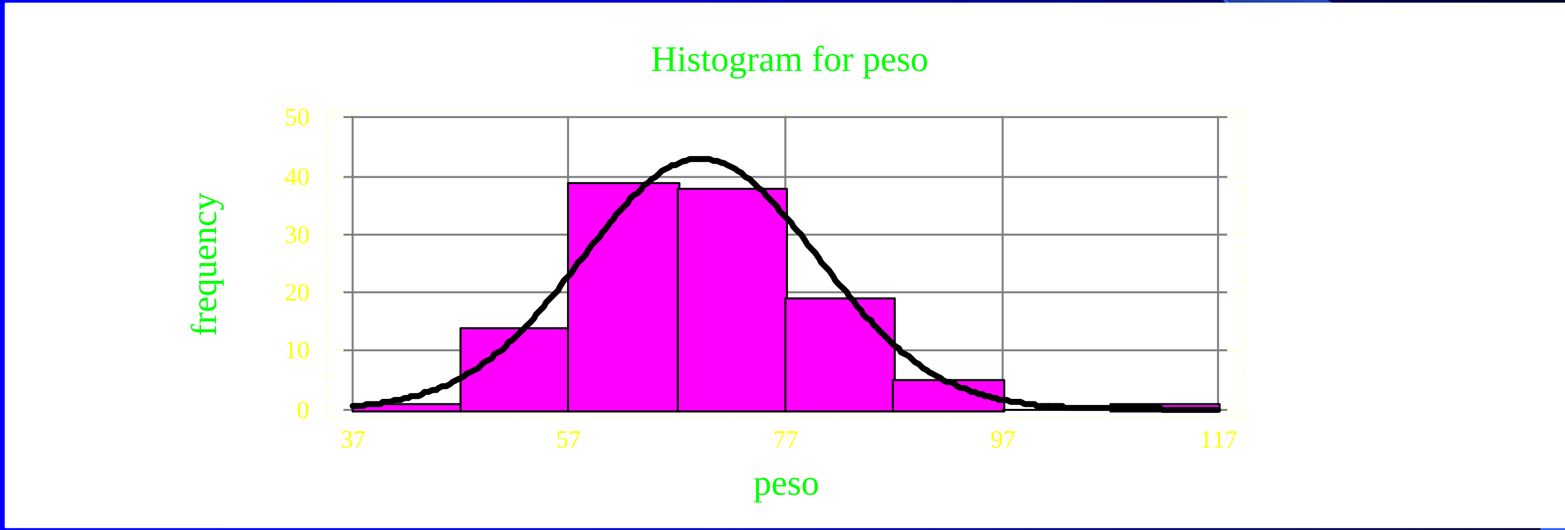
Nuestros datos serán las duraciones de estos cuatro componentes: t_1, t_2, t_3 y t_4.

Funciones asociadas al ADS

- En otros análisis estadísticos hemos utilizado la Función de Densidad y la Función de Distribución.
- En ADS además usaremos:
  - Función de Supervivencia o Función de Fiabilidad
  - Tasa de Fallos o Hazard Function

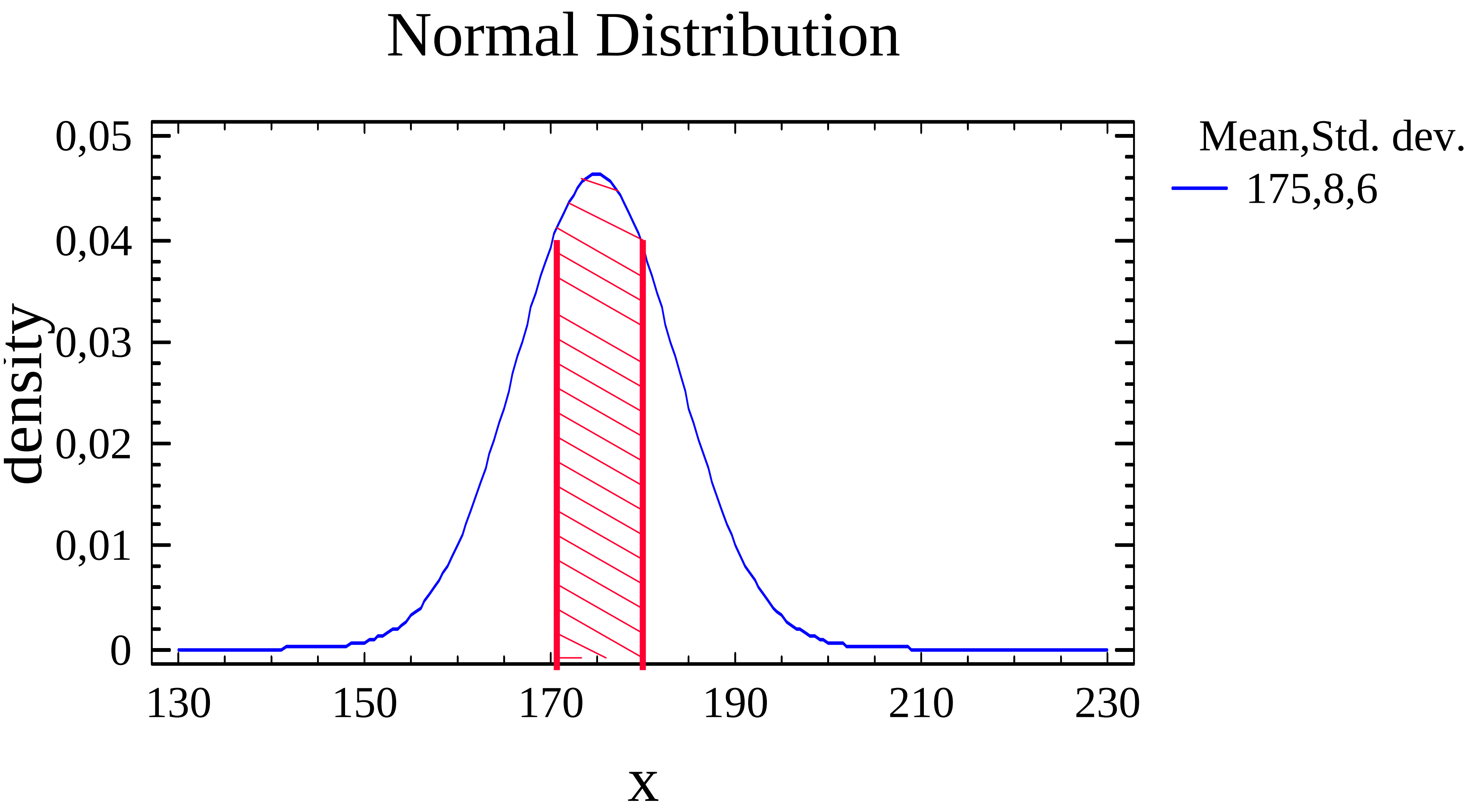
Función de Densidad (repaso)

Se denomina f(t)



El área comprendida bajo la función de densidad es la probabilidad de encontrar observaciones en ese intervalo.

Alturas de personas **N(175.8 ; 8.6)**



P(170<X<180)=Area bajo la curva. Lo calcula la máquina

Función de Distribución (repaso)

Se denomina F(t):

$$F(t) = \Pr(T \leq t) = \int_0^t f(x)dx.$$

Y proporciona la probabilidad de obtener valores menores de uno concreto.

Nuevas funciones

- Función de Supervivencia o de Fiabilidad
- Tasa de Fallos o Hazard Rate

Función de supervivencia

La probabilidad de que un individuo/componente sobreviva/funcione más allá de un instante *t*, viene dada por la función

$$S(t) = \Pr(T \geq t) = \int_t^{\infty} f(x)dx = 1 - F(t)$$

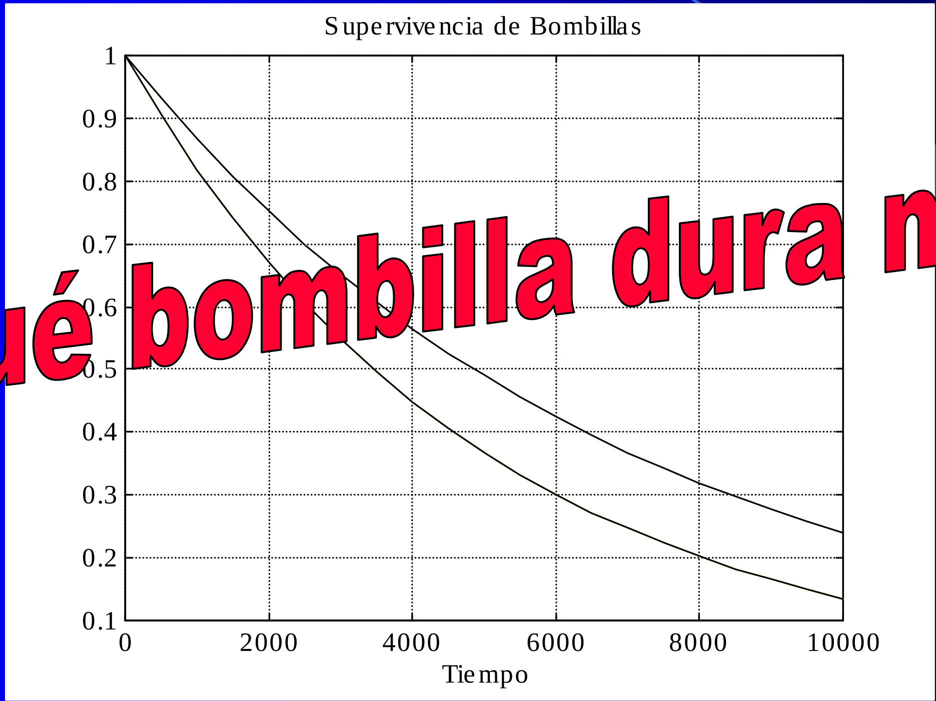
La función de supervivencia proporciona la probabilidad de que un componente esté funcionando al cabo de *t* horas.

Si un componente tiene una función de Fiabilidad:

S(1000)=0.89

quiere decir que la probabilidad de que el componente siga funcionando al cabo de 1000 horas es de 0.89.

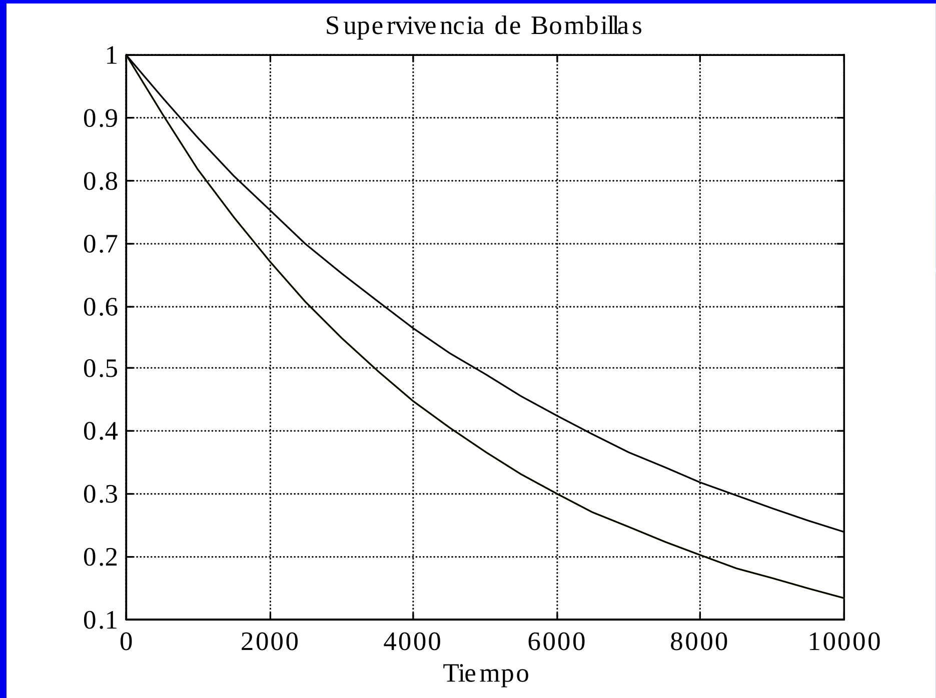
Función de supervivencia



¿Qué bombilla dura más?

La probabilidad de que ambas estén funcionando al cabo de 6000 horas es de 0.3 y 0.42 respectivamente.

Función de supervivencia



Evidentemente:

$$S(0) = 1$$

$$S(\infty) = 0$$

Tasa de Fallos

Para el análisis de procesos de duración, resulta especialmente indicada la hazard function -en fiabilidad se conoce como failure rate o tasa de fallo- que se define:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

Tasa de Fallos

Esta función proporciona la posibilidad de fallo inmediato dado que el componente está funcionando.

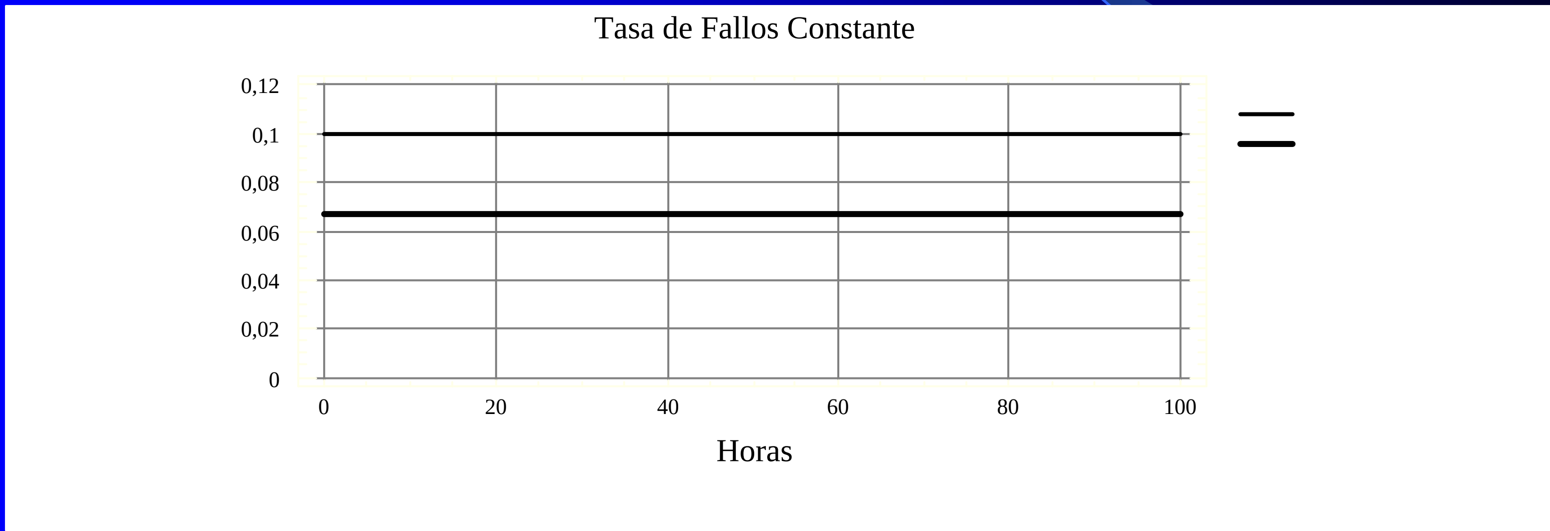
Es habitual encontrar funciones constantes, crecientes o decrecientes dependiendo del tipo de fenómeno estudiado. Los distintos procesos se van a definir según su tasa de fallos sea:

- **Creciente (IFR o Increasing Failure Rate)**
- **Decreciente (DFR o Decreasing Failure Rate)**
- **Constante (CFR)**

Tasa de fallos constante

- **Indica que la probabilidad de fallo instantáneo es la misma en cualquier momento y consecuentemente** el proceso no tiene memoria**, ya que la posibilidad de fallo estando funcionando, es idéntica en cualquier momento de la vida del componente**

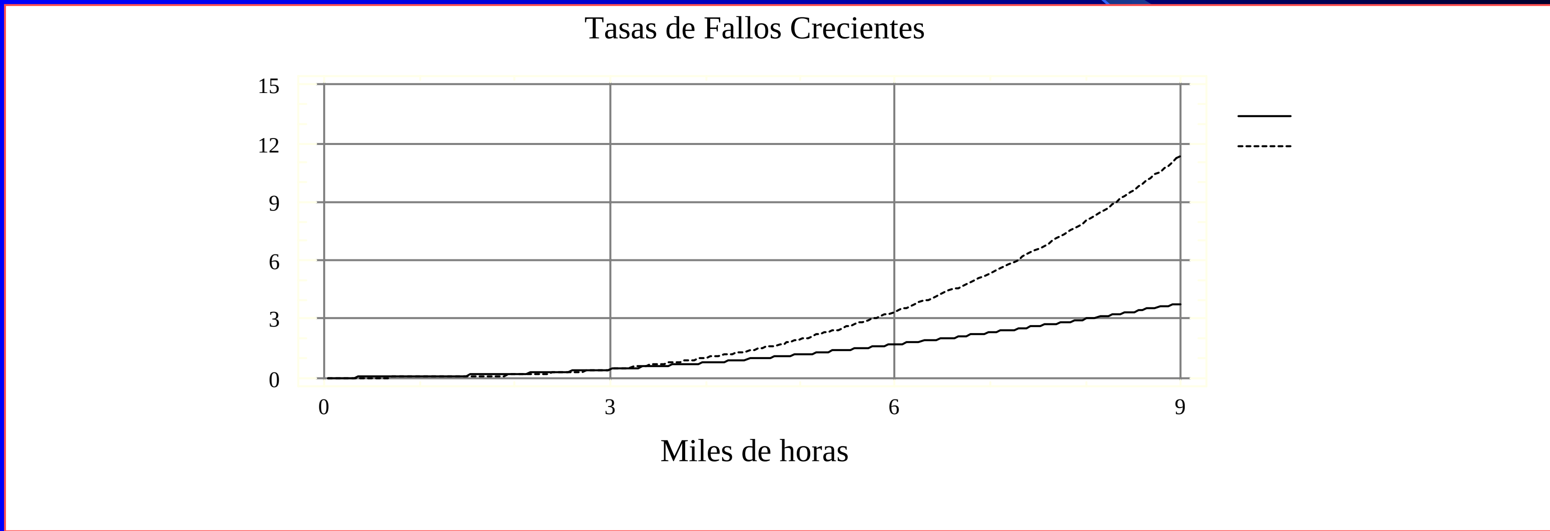
Tasa de fallos constante



Tasa de fallos creciente

- Surge, en la mayoría de los casos por desgastes y fatigas, es decir por un proceso de envejecimiento. La tasa de fallos creciente indica que la probabilidad de fallo inmediato, teniendo en cuenta que el componente está funcionando, se incrementa a medida que pasa el tiempo
- Evidentemente a medida que un componente se hace más viejo, su tasa de fallos tender á a crecer.

Tasa de fallos creciente

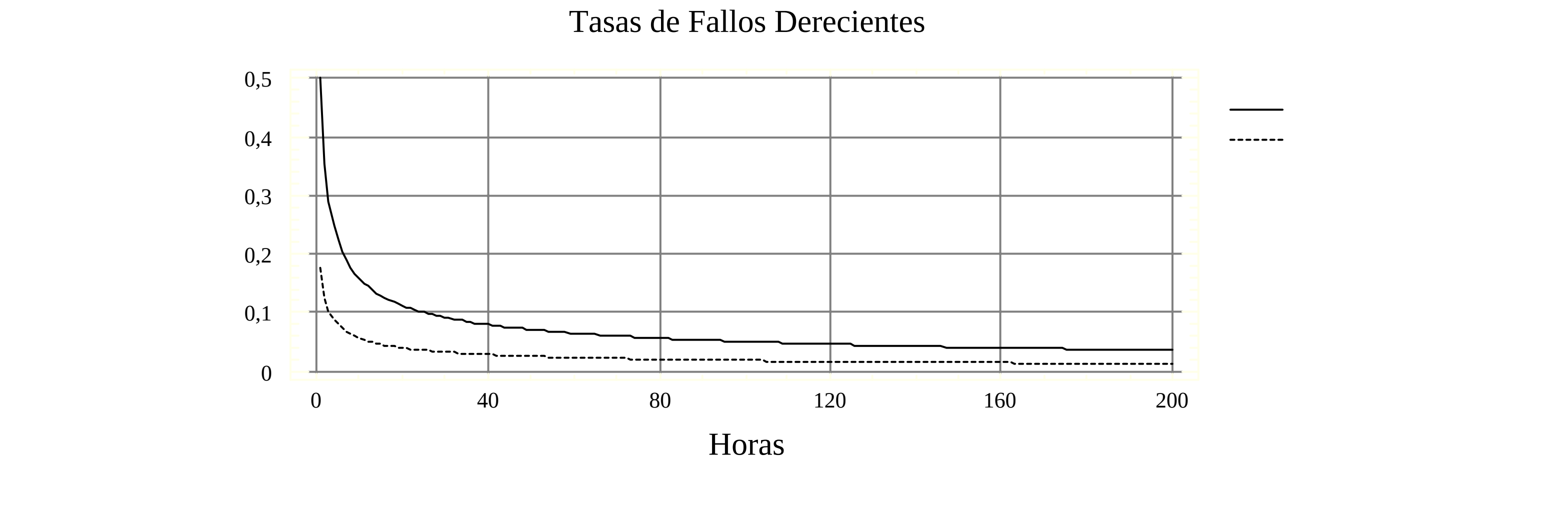


Tasa de fallos decreciente

Se observa en productos cuya probabilidad de fallo es menor cuando aumenta el tiempo de supervivencia.
Ésto aparece a menudo en cualquier tipo de materiales:

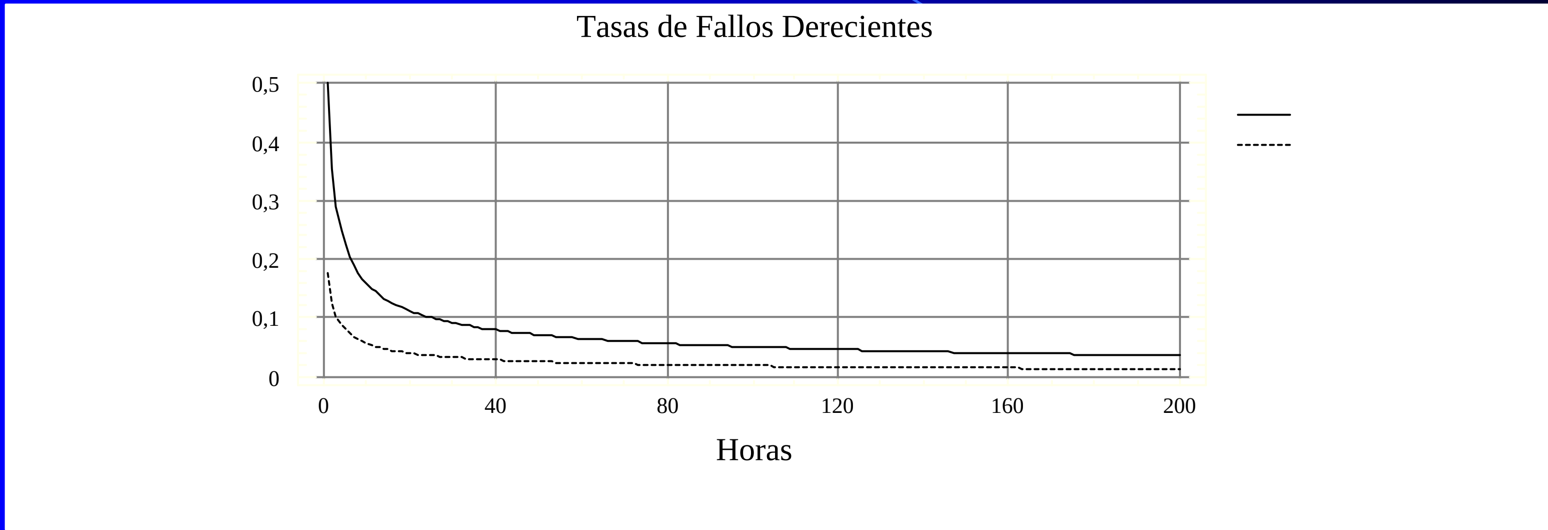
al principio de su funcionamiento la probabilidad de fallo es alta debido a la existencia de posibles defectos ocultos

Tasa de fallos decreciente



Si no fallan en las primeras 80 horas, la posibilidad de fallo se reduce notablemente en ambos casos.
El ensayo bajo stress permitir eliminar aquellos componentes que fallen al principio. De esta manera la empresa evita introducir en el mercado piezas defectuosas.

Tasa de fallos decreciente



La tasa de fallos decreciente aparece muy a menudo en estudios clínicos de supervivencia a intervenciones quirúrgicas:

El riesgo disminuye a medida que transcurre el postoperatorio.

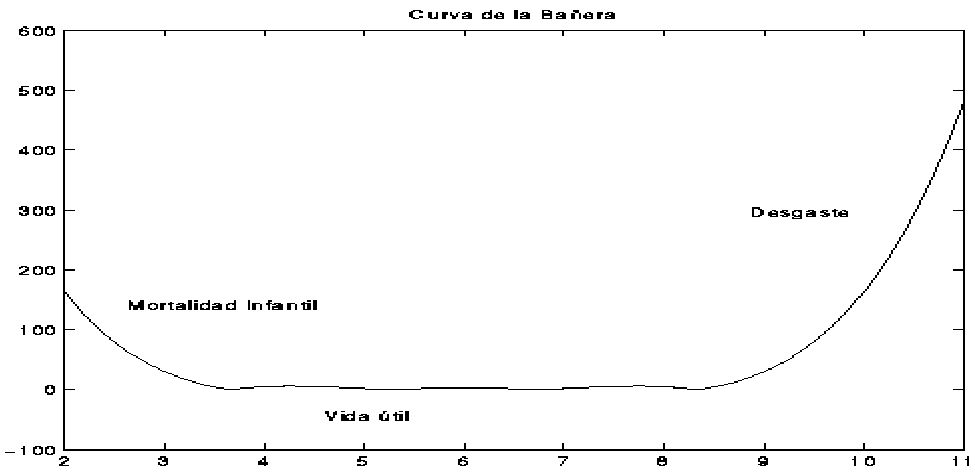
¿Cómo será la tasa de fallos de la vida humana?

HACEDLO

Curva de la bañera

- Generalización de los procesos anteriores
- Muy común en la práctica
- un elemento que se comporta inicialmente de forma decreciente (a esta zona se le denomina de mortalidad infantil)
- en su vida media con una probabilidad de fallo casi constante (zona de vida útil)
- finalmente con probabilidad de fallo que aumenta con la edad (zona de deshecho, wearout)

Curva de la Bañera

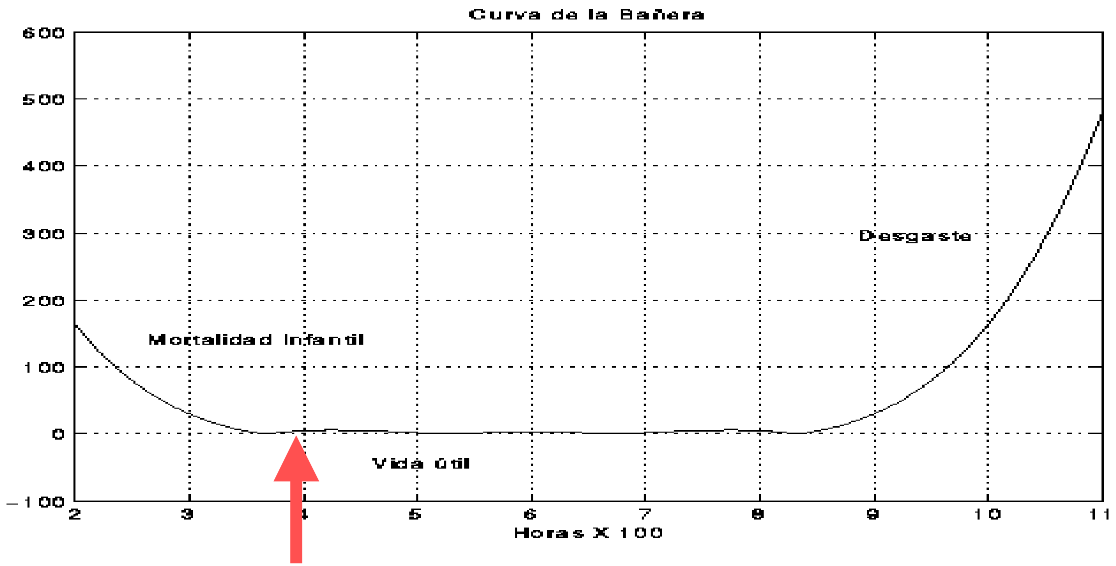


Cuando la tasa de fallo del elemento responde a la curva de la bañera es conveniente realizar un ensayo acelerado del mismo (en condiciones de stress) para que supere la zona de mortalidad infantil o de Burn-in.

Periodos de Garantía y ensayos acelerados

- Producto con tasa de fallos con mortalidad infantil (DFR o curva de la bañera) la empresa se enfrenta a un problema:
  - Sus productos tienen mayor posibilidad de fallo en los primeros momentos de funcionamiento debido a la existencia de defectos ocultos.
  - Sin embargo, la empresa no puede detectar fácilmente esos fallos.
- Posibilidad interesante:
  - determinar cuando comienza la vida útil del producto y ofrecer a los clientes una garantía de funcionamiento durante ese periodo de funcionamiento problemático.
  - Una vez superado el periodo crítico, la empresa está razonablemente segura de que el producto tiene una posibilidad de fallos reducida

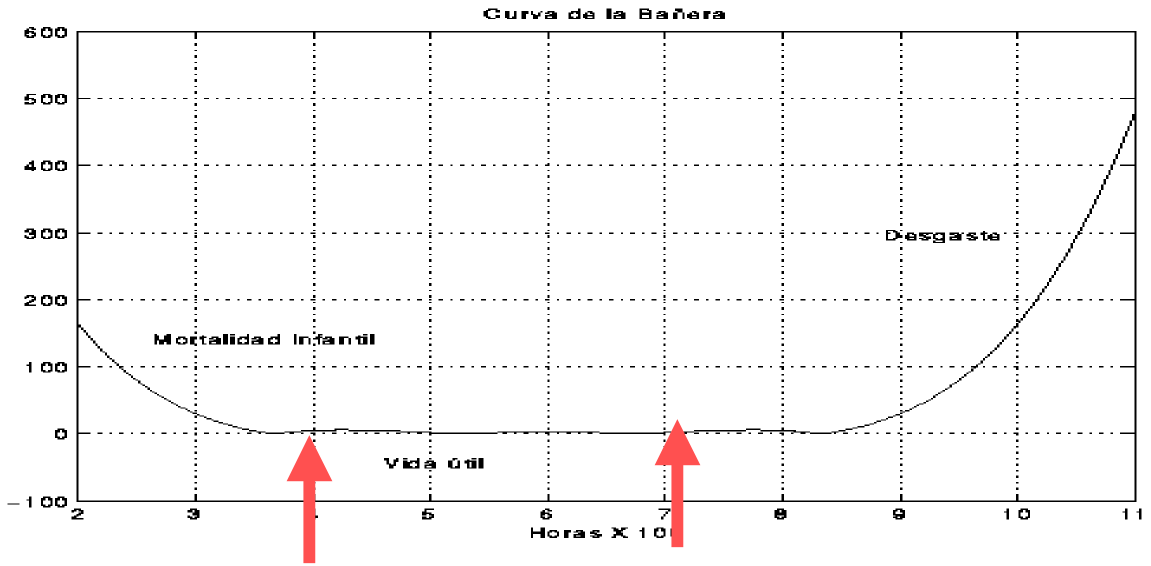
Periodos de Garantía y ensayos acelerados



En el ejemplo, la empresa garantizaría el producto durante, al menos, 400 horas.

Algunas empresas están desarrollando estrategias comerciales basadas en ampliar el periodo de garantía a la vida útil del producto.

Un producto tiene una tasa de fallos muy baja durante su vida útil. Entonces, el es muy probable que el producto empiece a fallar cuando alcance la zona de desgaste. Si esto es así, la empresa puede prolongar a muy bajo coste la garantía incluyendo una importante parte de la zona útil del producto, resaltando que el producto es muy fiable.



La empresa podría incrementar la garantía hasta 700 horas con un coste adicional muy bajo

Estrategia para coches, electrodomésticos......

Algunos productos, sin embargo no pueden fallar:

- Componentes clave de determinados procesos como por ejemplo válvulas de centrales nucleares, aviones, mecanismos de seguridad, etc, no pueden tener problemas en los primeros momentos de su aplicación debido a la tasa de fallos decreciente.

Una posibilidad en estos casos:

Probar el componente sometido a condiciones limite. Por ejemplo, si una válvula en una central nuclear debe funcionar a 10 atmósferas de presión y 100ºC de temperatura, se somete las válvulas a un ensayo de funcionamiento a 30 atmósferas y 200ºC.

Los defectos ocultos que provocan la mortalidad infantil afloran y la fiabilidad del aparato aumenta.

Las pruebas aceleradas o bajo stress se realizan únicamente en sistemas que requieren una alta fiabilidad desde el principio. En otras condiciones no suele ser rentable

Modelos utilizados en Fiabilidad. Datos Completos

- Ajustaremos modelos de probabilidad para poder generalizar los conocimientos que tenemos a partir de una pequeña muestra de componentes.
- El criterio de elección de un modelo se basará en técnicas descriptivas y especialmente en el conocimiento teórico que tengamos del proceso.
- Este conocimiento nos permitirá saber en muchas ocasiones que el proceso tiene tasa de fallos creciente, decreciente o en forma de bañera.

Modelo exponencial

- **El modelo exponencial es bien conocido. Su función de densidad es:**

$$f(t)=\frac{1}{\theta}\exp(-t/\theta)$$

Supervivencia: $S(t)=\exp(-t/\theta)$

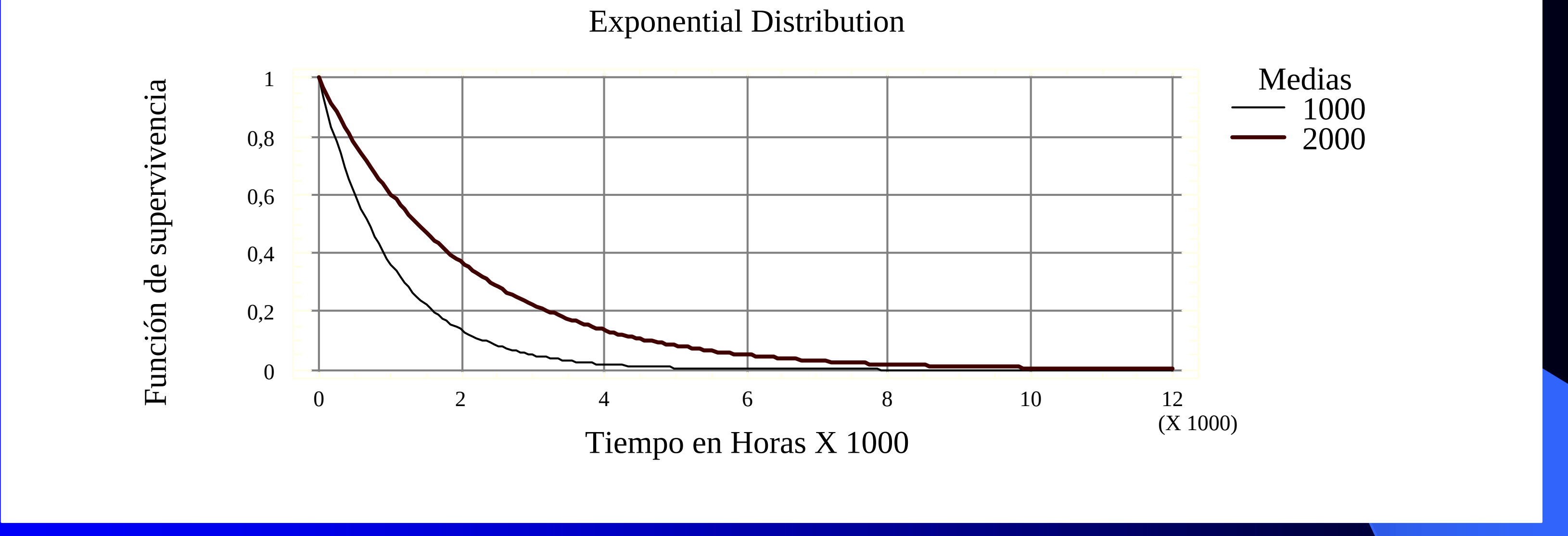
Tasa de fallos: $h(t)=1/\theta$

Esperanza: $E(t)=\theta$

El modelo exponencial

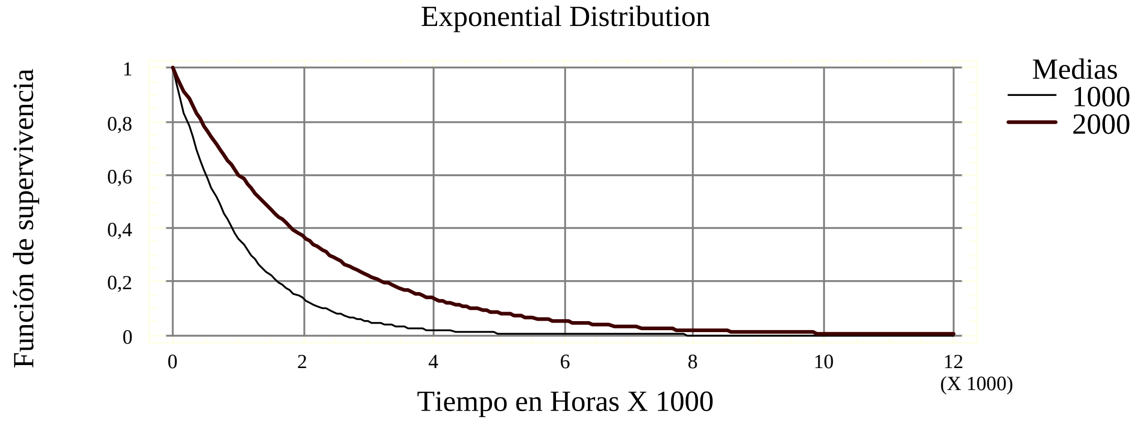
- Es el único con tasa de fallos constante.
- La probabilidad de fallar condicionada a que el elemento este en uso no varia con el tiempo. Esta propiedad se denomina falta de memoria.

Función de Supervivencia



Si son bombillas ¿Cuál es mejor?

Función de Supervivencia

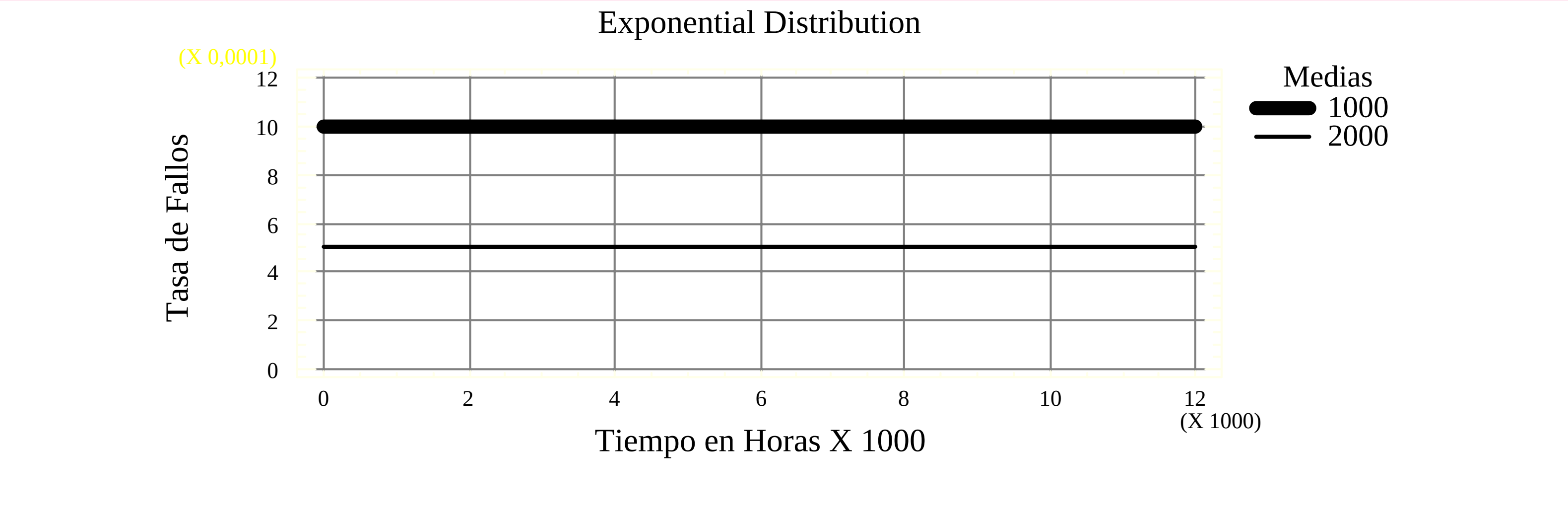


E(una)=1000 horas y E(otra)=2000 horas

La probabilidad de que el componente con vida media de 1000 horas funcione mas de 2000 horas es del 13.5%.
Para el componente de 2000 horas de duración media es de 36.7%.
Estas cifras se obtienen de la función de supervivencia

$$S(t) = e^{-t/\theta} = e^{-2000/1000} = 0.135$$

Tasa de Fallos



Si son bombillas ¿Cuál es mejor?

Modelo Weibull

- **El modelo Weibull tiene la siguiente función de densidad:**

$$f(t) = \lambda\beta(\lambda t)^{\beta-1}\exp(-(\lambda t)^{\beta}) \qquad t \geq 0$$

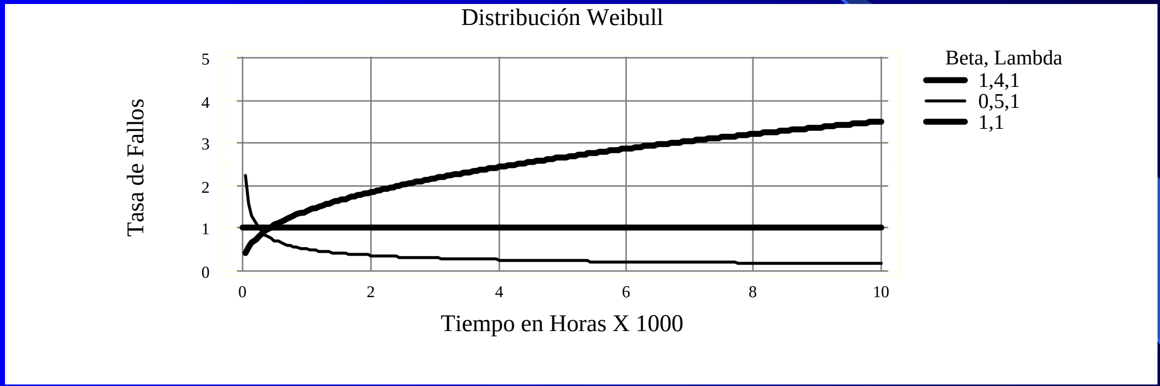
$$S(t) = \exp(-(\lambda t)^{\beta}) \qquad t \geq 0$$

$$h(t) = \lambda\beta(\lambda t)^{\beta-1} \qquad t \geq 0$$

Modelo Weibull

- Tasa de fallos: $$h(t) = \lambda\beta(\lambda t)^{\beta-1}$$
- Según sean los valores de beta puede presentar tasas de fallo crecientes, decrecientes o constantes.
- Cuando beta=1 el modelo Weibull se convierte en exponencial y presenta tasa de fallos constante. El modelo exponencial es por tanto un caso particular del modelo Weibull.
- Cuando beta>1 el modelo presenta tasa de fallos creciente.
- Cuando beta<1 el modelo presenta tasa de fallos decreciente.

Modelo Weibull: tasas de fallos



El modelo Weibull es muy versátil y en la practica es uno de los mas utilizados

Estimación Paramétrica

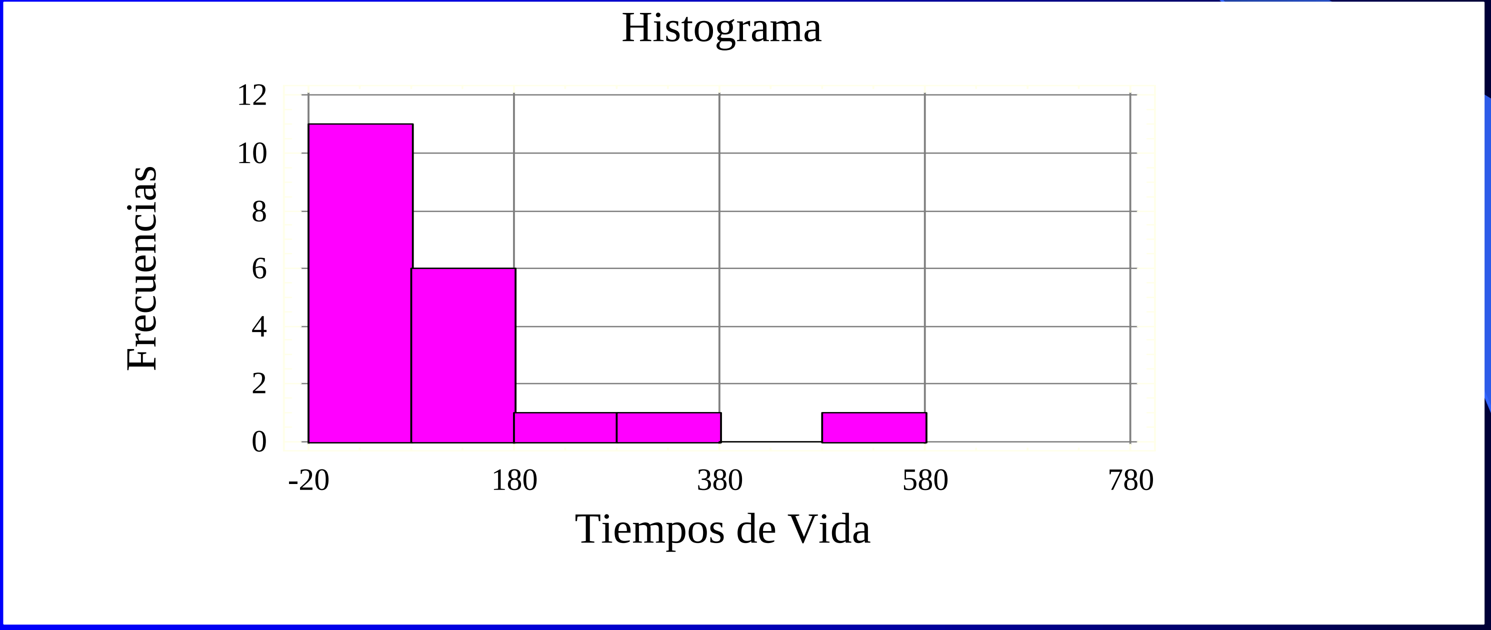
- El proceso de ajuste de modelos estadísticos a partir de datos muestrales es simple:
  - Se estudian los datos mediante técnicas de estadística descriptiva
  - Se elige un modelo de distribución de probabilidad
  - Se estima
  - Se realiza una diagnosis para detectar posibles errores.

Ejemplo 1

- **Se ha realizado un ensayo para estudiar la duración de vida de unos componentes electrónicos. Para ello se han puesto 20 elementos a prueba y se han observado hasta el fallo. Los tiempos de vida recogidos han sido los siguientes:**
- **58,91 158,8 25,16 80,26 77,85 105,4 95,97 87,29 81,49 16,39 79,10 36,89 68,05 21,31 209,41 519,26 34,24 44,33 283,2 8,33**

Ejemplo 1

- 58,91 158,8 25,16 80,26 77,85 105,4 95,97 87,29 81,49 16,39 79,10 36,89 68,05 21,31 209,41 519,26 34,24 44,33 283,2 8,33



Ajuste del modelo exponencial

- El modelo exponencial puede ser adecuado para estos datos. Optaremos por una distribución exponencial con:

$$f(t;\theta) = 1/\theta \exp(-t/\theta).$$

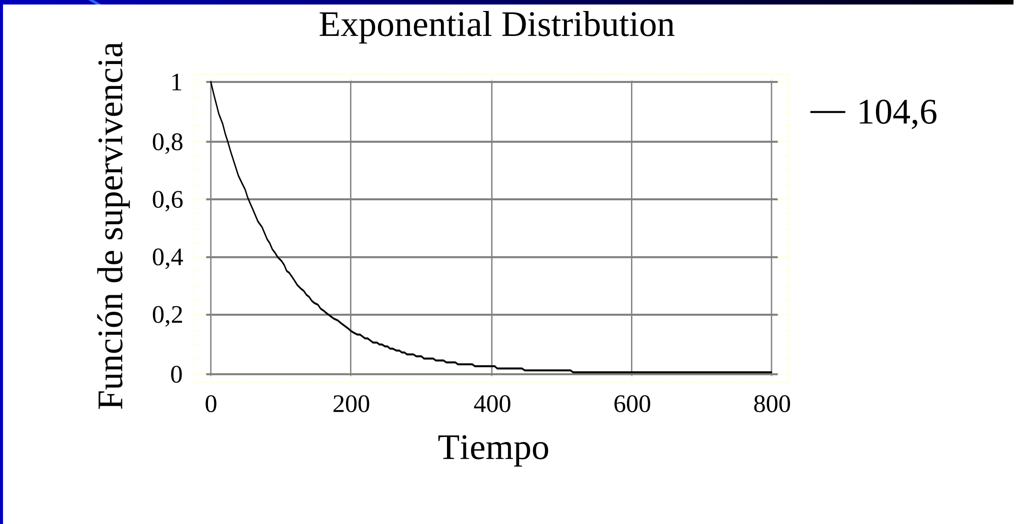
$$\widehat{\theta} = \frac{\sum_{i=1}^{n} t_i}{n} = \bar{t}$$

- En nuestro caso

Theta=media de tiempos=104.6

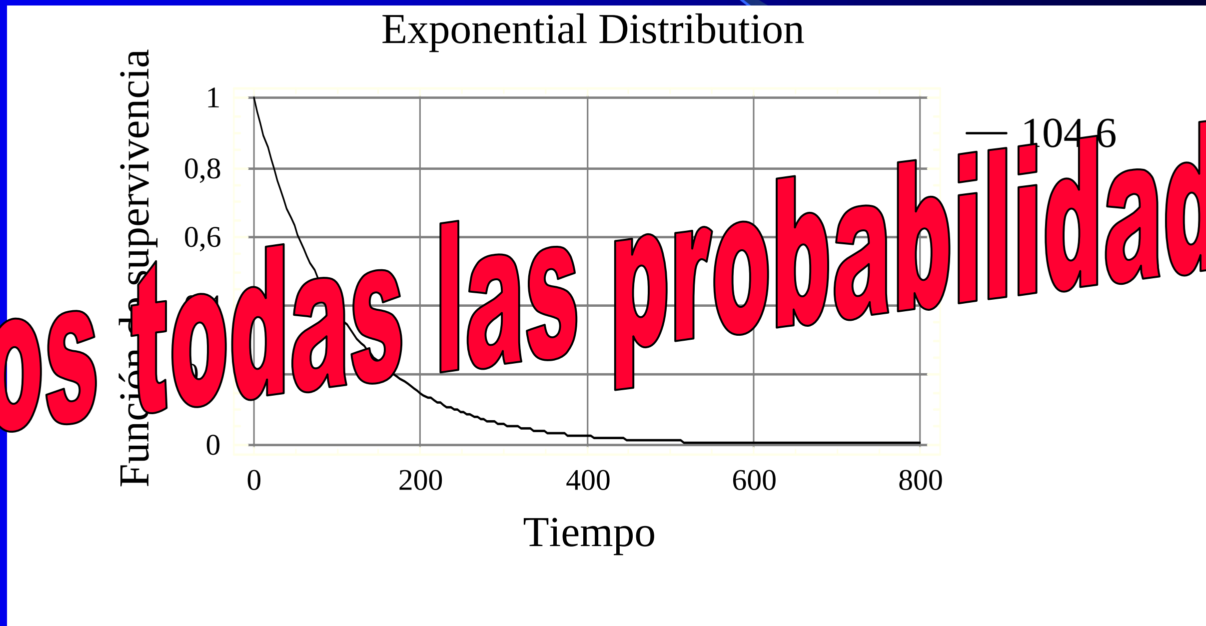
Ajuste del modelo exponencial

- **A partir de aquí podemos inferir muchas propiedades de nuestro componente.**
- **Por ejemplo, la probabilidad de que un componente dure mas de 200 horas será:**



$$\Pr(T > 200) = S(200) = e^{-0.0096.200} = 0.1466$$

Ajuste del modelo exponencial



Conocemos todas las probabilidades

Métodos gráficos para elegir el modelo adecuado.

- A mano:
  - Estimar la funcion de distribucion empirica de los datos y representarla en unas escalas tales que si el modelo elegido es correcto los datos presenten aspecto lineal.
- En ordenador:
  - Lo normal. Hace lo mismo pero de forma mecánica.

A mano

- Tenemos los datos de tiempos de fallo de una serie de componentes.
- 1. Ordenamos de menor a mayor
- 2. Estimación de la Función de Distribución corregida mediante:

$$F_i=(i-0.3)/(n+0.4)$$

- Construcción del gráfico adecuado hasta que los datos formen una recta.
- Estimación de los parámetros

Construcción del gráfico exponencial

- **Si los datos son exponenciales, la función de supervivencia será:**

$$S(t)=e^{-t/\theta}$$

- **Tomando logs**

$$\text{Log } S(t)=-t/\theta$$

- **Como $F(t)=1-S(t)$**

$$-\text{Log } (1-F(t))= t/\theta$$

- **Por tanto si en un gráfico se coloca en el eje Y la variable $Y= -\text{Log } (1-F(t))$ y en el eje X la variable t**
- ***Si los datos son exponenciales deberían estar alineados***

A mano: Ejemplo

1. Datos ordenados de menos a mayor

3,04 4,45 6,25 37,1 42,7 76,6 76,7 103,9 107,7 110,8 114,6
121,2 130,2 220 236,8 245,6 314,8 407,9 499 627,4.

2. Estimación de la Función de Distribución corregida mediante:

$$F_i=(i-0.3)/(n+0.4)$$

A mano: Ejemplo

3,04 4,45 6,25 37,1 42,7 76,6 76,7 103,9 107,7 110,8 114,6 121,2 130,2 220 236,8 245,6 314,8 407,9 499 627,4

2. Estimación de la Función de Distribución corregida mediante:

$$F_i=(i-0.3)/(n+0.4)$$

Tiempos	Orden;	$F = \frac{i-0.3}{20-0.4}$
3.04	1	0.03
4.45	2	0.08
6.25	3	0.13
37.10	4	0.18
42.7	5	0.23
76.6	6	0.28
76.7	7	0.33
103.9	8	0.38
107.7	9	0.43
110.8	10	0.48
114.6	11	0.52
121.2	12	0.57
130.2	13	0.62
220	14	0.67
236.8	15	0.72
245.6	16	0.77
314.8	17	0.82
407.9	18	0.87
499.2	19	0.92
627.4	20	0.97

A mano: Ejemplo

En un gráfico se coloca en el eje Y la variable

Y= -Log (1-F(t))

y en el eje X la variable t

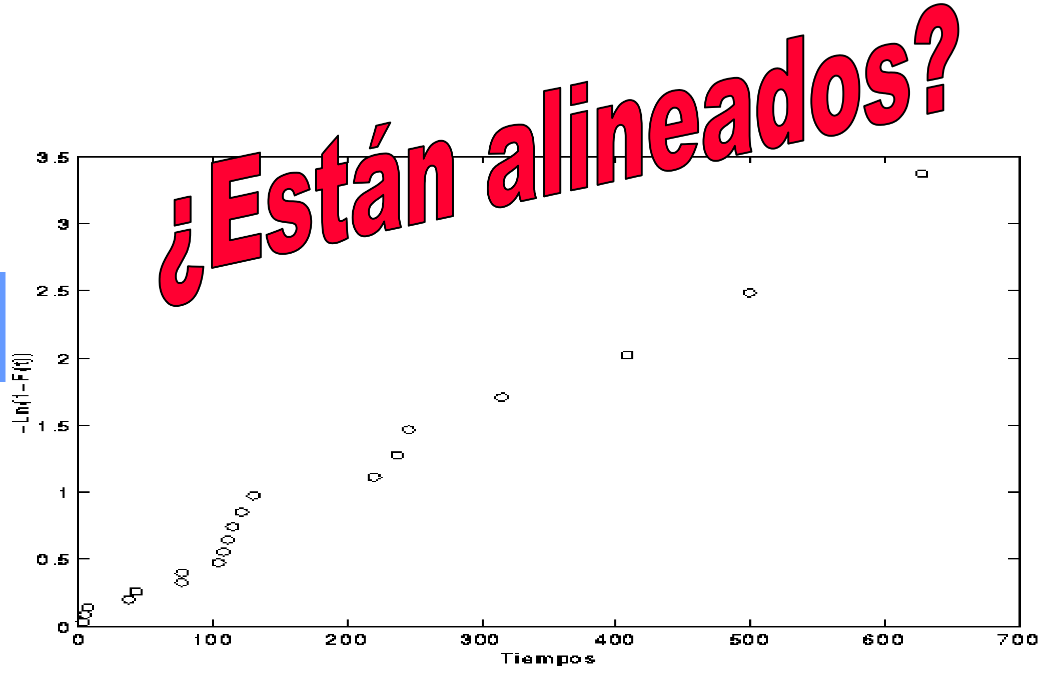
Tiempos	Orden; $F = \frac{i-0.3}{20-0.4}$		$-\ln(1 - F(T))$
3.04	1	0.03	.03
4.45	2	0.08	.08
6.25	3	0.13	.14
37.10	4	0.18	.20
42.7	5	0.23	.26
76.6	6	0.28	.32
76.7	7	0.33	.39
103.9	8	0.38	.47
107.7	9	0.43	.55
110.8	10	0.48	.64
114.6	11	0.52	.74
121.2	12	0.57	.85
130.2	13	0.62	.97
220	14	0.67	1.11
236.8	15	0.72	1.27
245.6	16	0.77	1.46
314.8	17	0.82	1.70
407.9	18	0.87	2.02
499.2	19	0.92	2.44
627.4	20	0.97	3.38

A mano: Ejemplo

En un gráfico se coloca en el eje Y la variable Y= -Log (1-F(t))

y en el eje X la variable t

-Log (1-F(t))



t

Alineación de datos para el modelos Weibull

El modelo Weibull tiene una funcion de supervivencia:

$$S(t) = \exp(-(\lambda t)^{\beta})$$

Tomando Logaritmos

$$\ln S(t) = -(\lambda t)^{\beta}$$

$$\ln(1 - F(t)) = -(\lambda t)^{\beta}$$

Alineación de datos para el modelos Weibull

$$\ln(1 - F(t)) = -(\lambda t)^{\beta}$$

volviendo a tomar logaritmos:

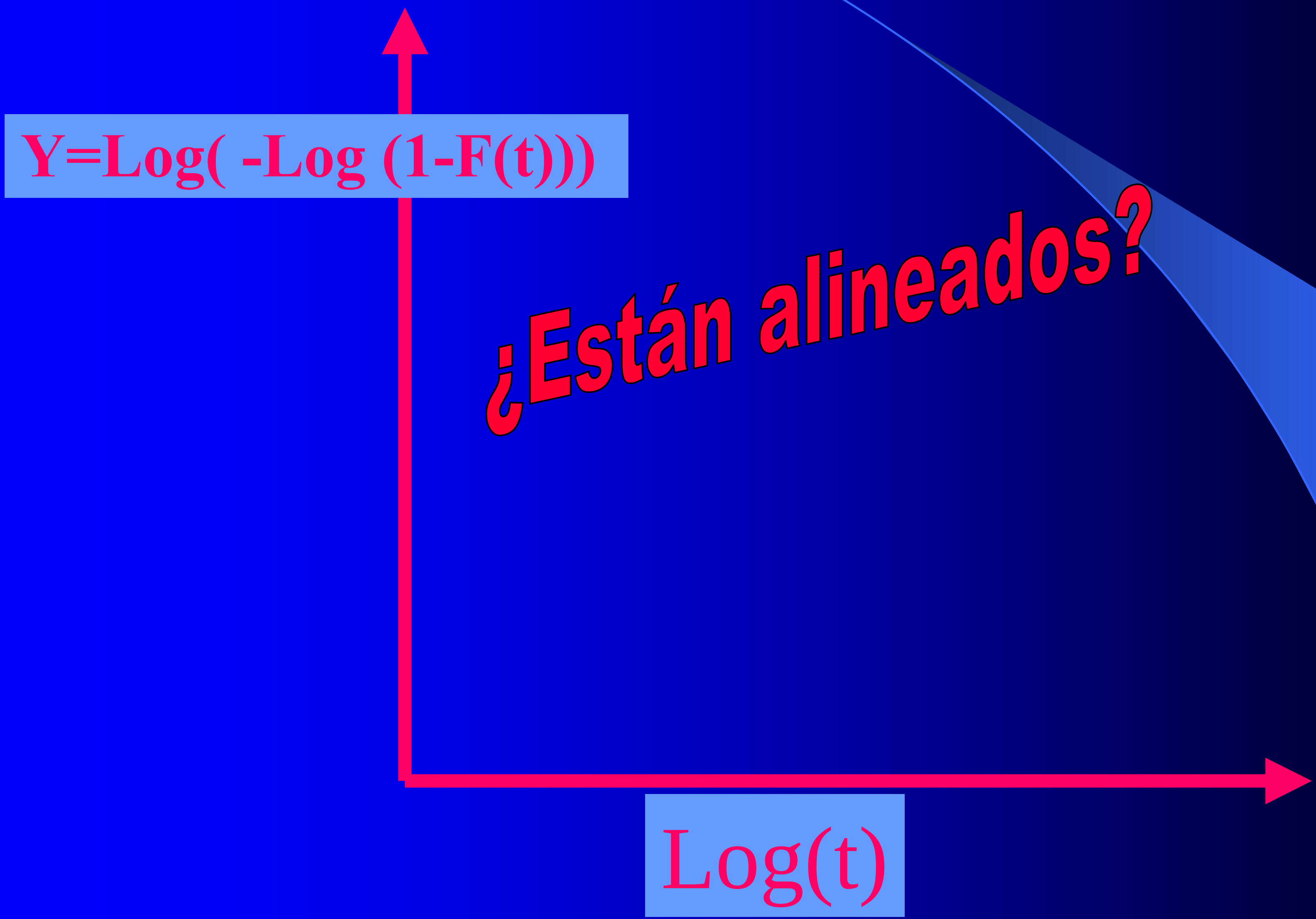
$$\ln(-\ln(1 - F(t))) = \beta \ln(\lambda t) = \beta \ln(\lambda) + \beta \ln(t)$$

$$\ln(-\ln(1 - F(t))) = \beta \ln(\lambda) + \beta \ln(t)$$

$$Y = \alpha + \beta X$$

Weibull

En un gráfico se coloca en el eje Y la variable Y=Log(-Log (1-F(t)))
y en el eje X la variable log(t)



Weibull ejemplo

Se tienen datos de la duracion de 20 componentes. Los datos son:

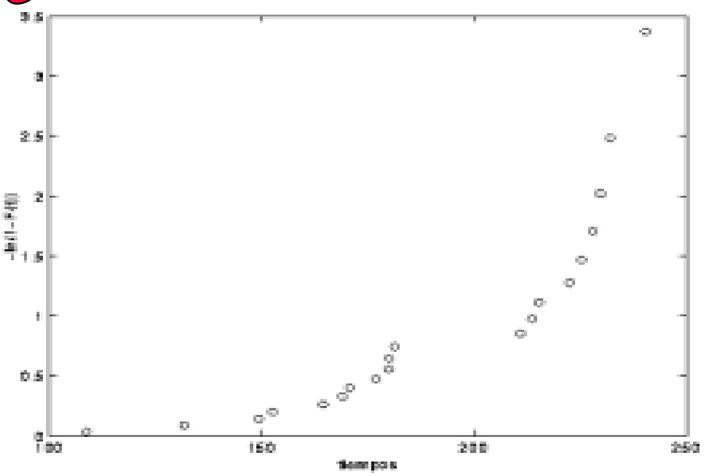
58,435 261,126 56,6706 230,788 183,028 19,3203 128,744 58,0366 131,247 397,636 79,4311 180,2 28,2613 131,948 323,421 219,182 167,721 130,961 207,719 285,59

Vamos a realizar un grafico exponencial para comprobar si los datos son exponenciales.

Weibull en gráfico exponencial

¿Están alineados?

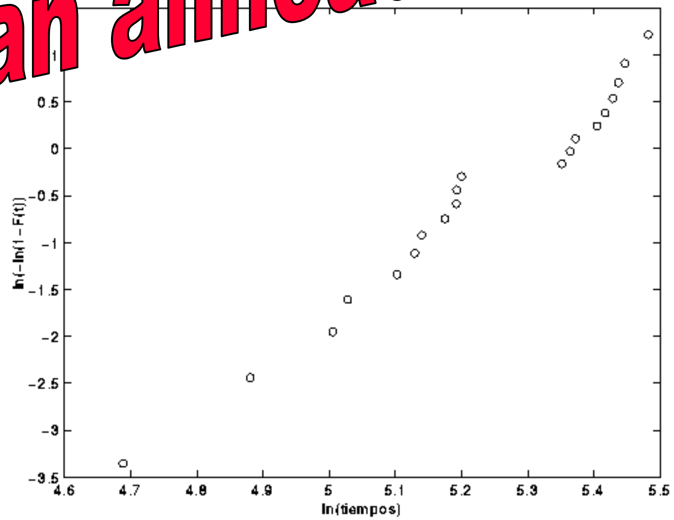
-Log (1-F(t))



t

Weibull ejemplo

¿Están alineados?



Y=Log(-Log (1-F(t)))

Log(t)

Gráficos en ordenador

- Elaborar estos gráficos es laborioso
- En la práctica se hacen con ordenador.
- Statraphics lo hace mucho mejor que nosotros
- Así que NO HAREMOS GRÁFICOS A MANO QUE ES HORRIBLE

Gráficos en ordenador

- Se introducen los datos
- Se va a DESCRIBE y Distribution Fitting
- Se va a Weibull analysis
- Se escogen los datos adecuados y se pide el Weibull Plot.

Gráficos en ordenador

- El resultado es:

Estimación Weibull

- Una vez elegido el modelo lo estimamos
- Método de máxima verosimilitud
- En el caso exponencial ya se ha visto
- Θ=media de los tiempos
- El caso Weibull es más complejo ecuaciones que hay que resolver por métodos numéricos.
- Lo estima el ordenador

Estimación Weibull

- **El ordenador estima λ y β.**
- **Los datos eran Weibull como se vio en el gráfico.**

Se tienen datos de la duracion de 20 componentes. Los datos son:

58,435 261,126 56,6706 230,788 183,028 19,3203 128,744 58,0366 131,247 397,636 79,4311 180,2 28,2613 131,948 323,421 219,182 167,721 130,961 207,719 285,59

Estimación Weibull

$$\hat{\lambda} = 203.78$$

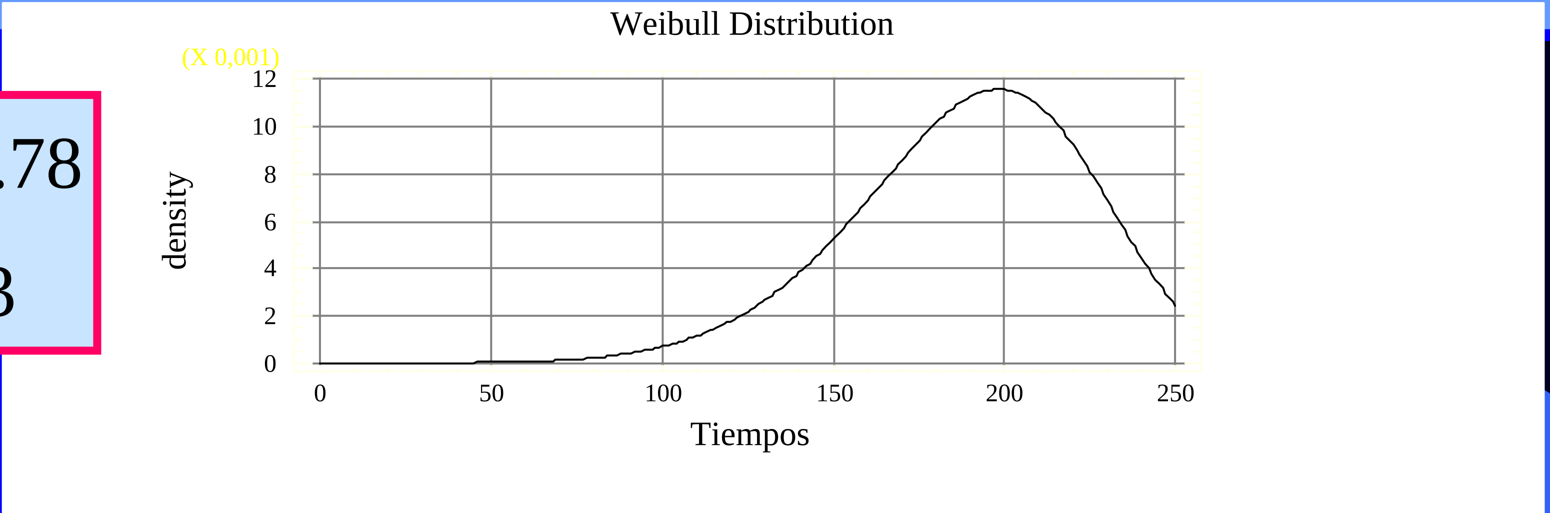
$$\hat{\beta} = 6.33$$

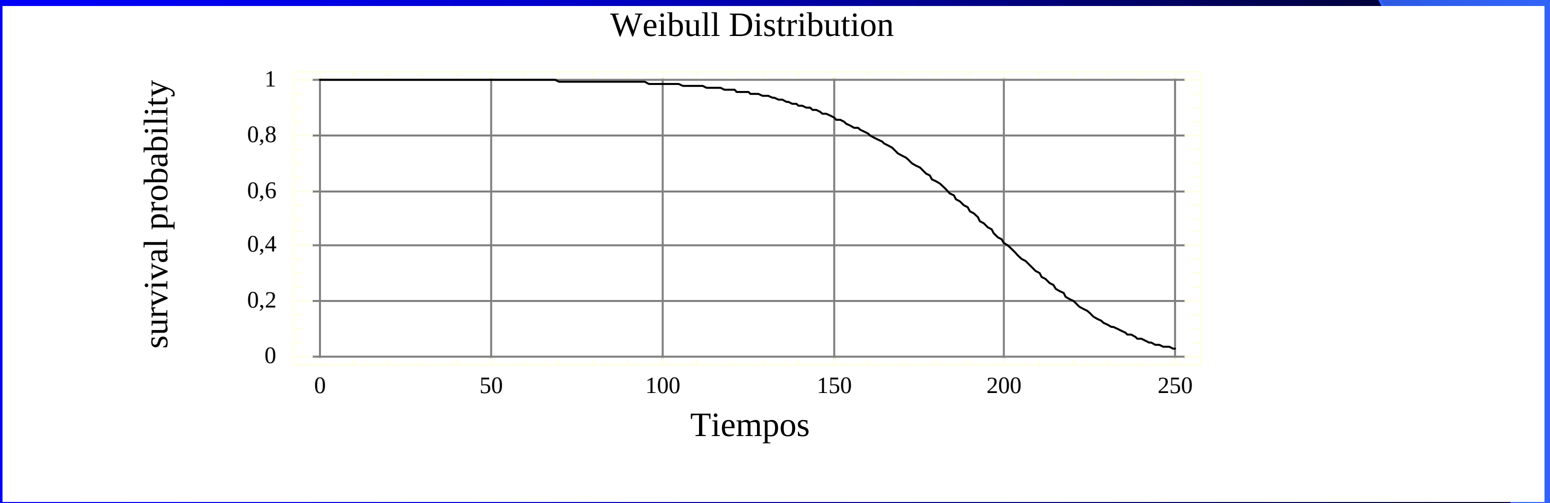
Con estos valores es posible conocer muchas cosas de nuestro componente.
Por ejemplo la probabilidad de que falle antes de 100 horas es de 0.11. Y la de que falle antes de 250 horas es de 0.97

Estimación Weibull

$$\hat{\lambda} = 203.78$$

$$\hat{\beta} = 6.33$$





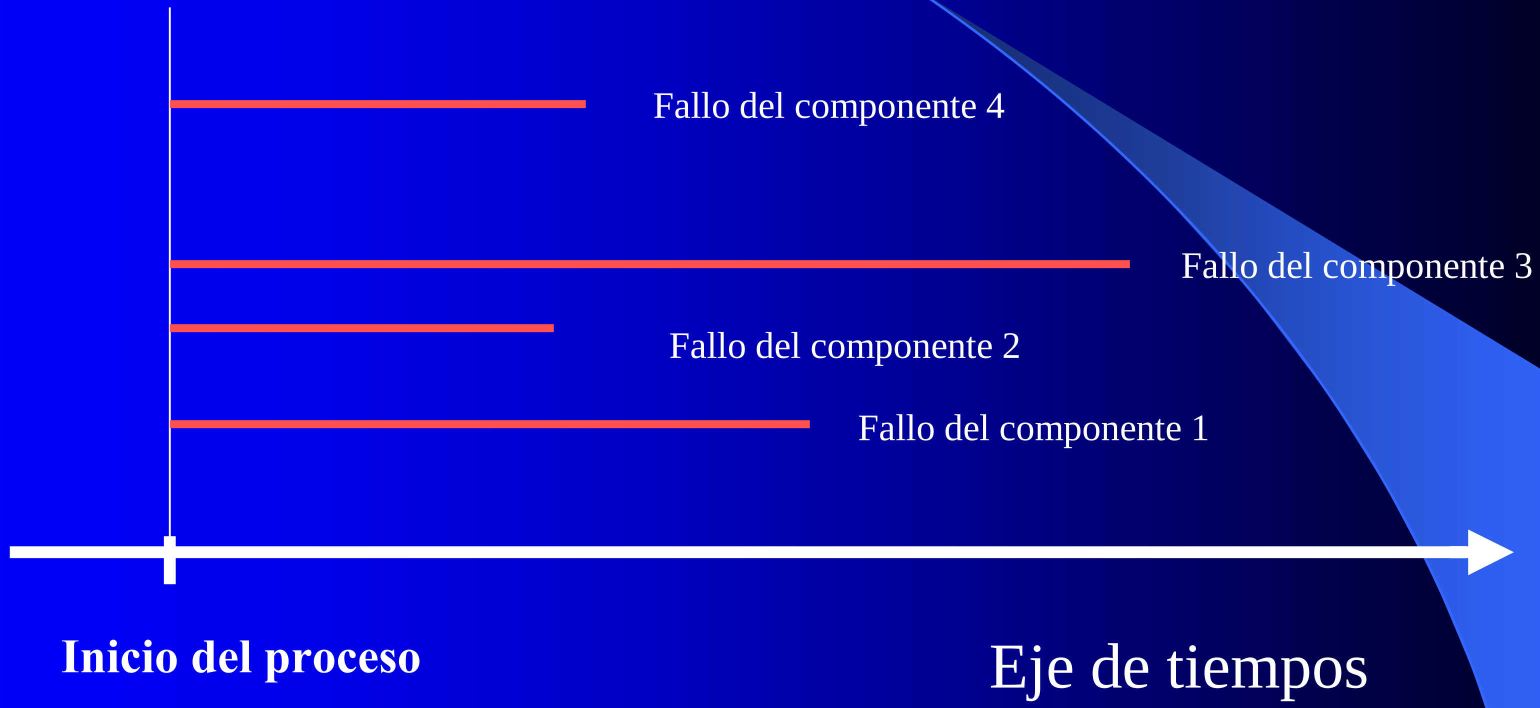
Datos incompletos

Censura

Censura

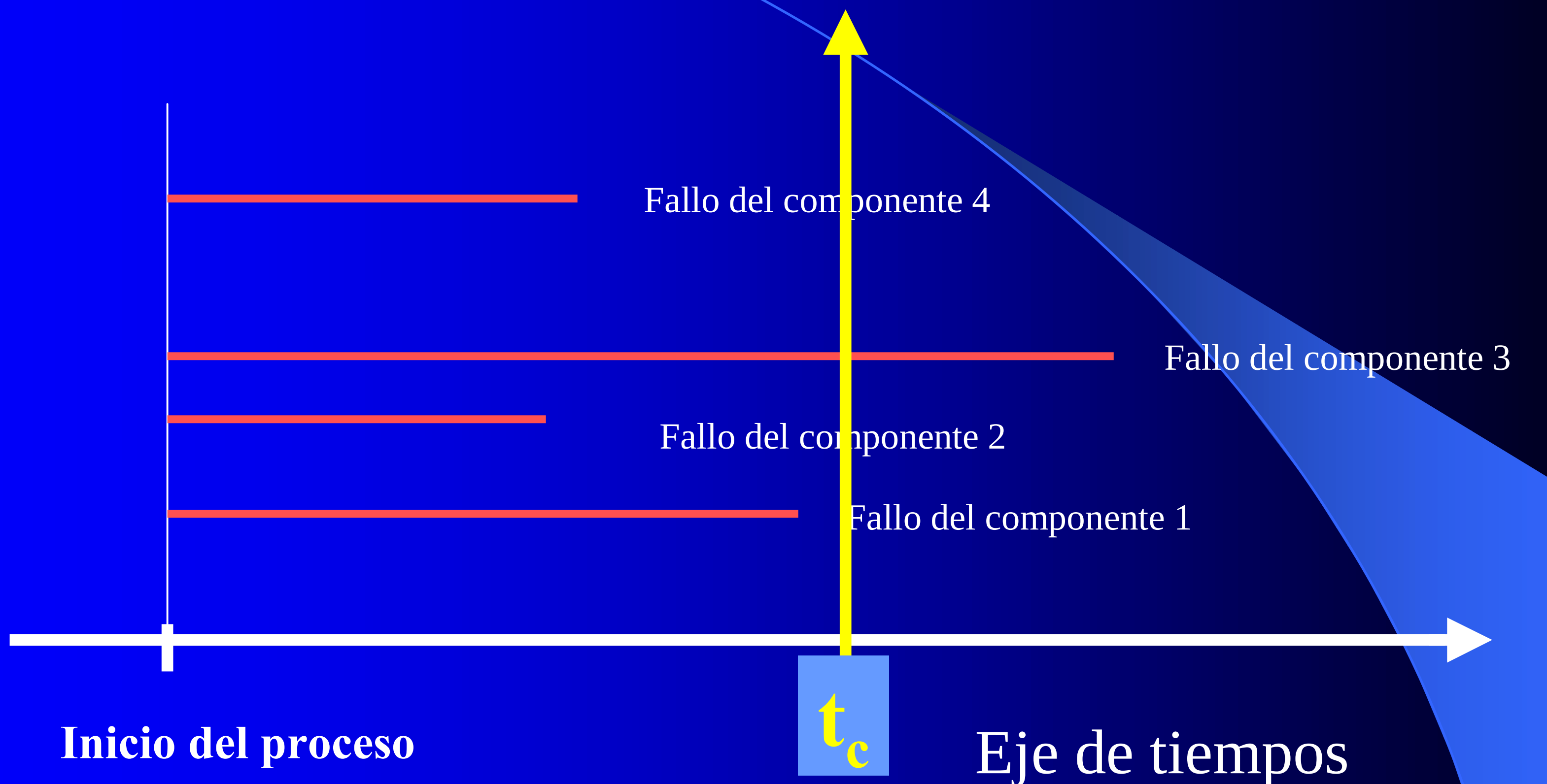
- **Una observación esta censurada cuando solo contiene información parcial sobre la variable a estudiar.**
- **Esta situación es muy frecuente: la longitud del intervalo entre tránsitos impide muchas veces el seguimiento de la muestra hasta el transito final.**
- **Hay tres tipos de censura:**
  - **Censura por la derecha**
  - **Censura por la izquierda**
  - **Censura por intervalos**

Sin censura



Nuestros datos serán las duraciones de estos cuatro componentes: t_1, t_2, t_3 y t_4.

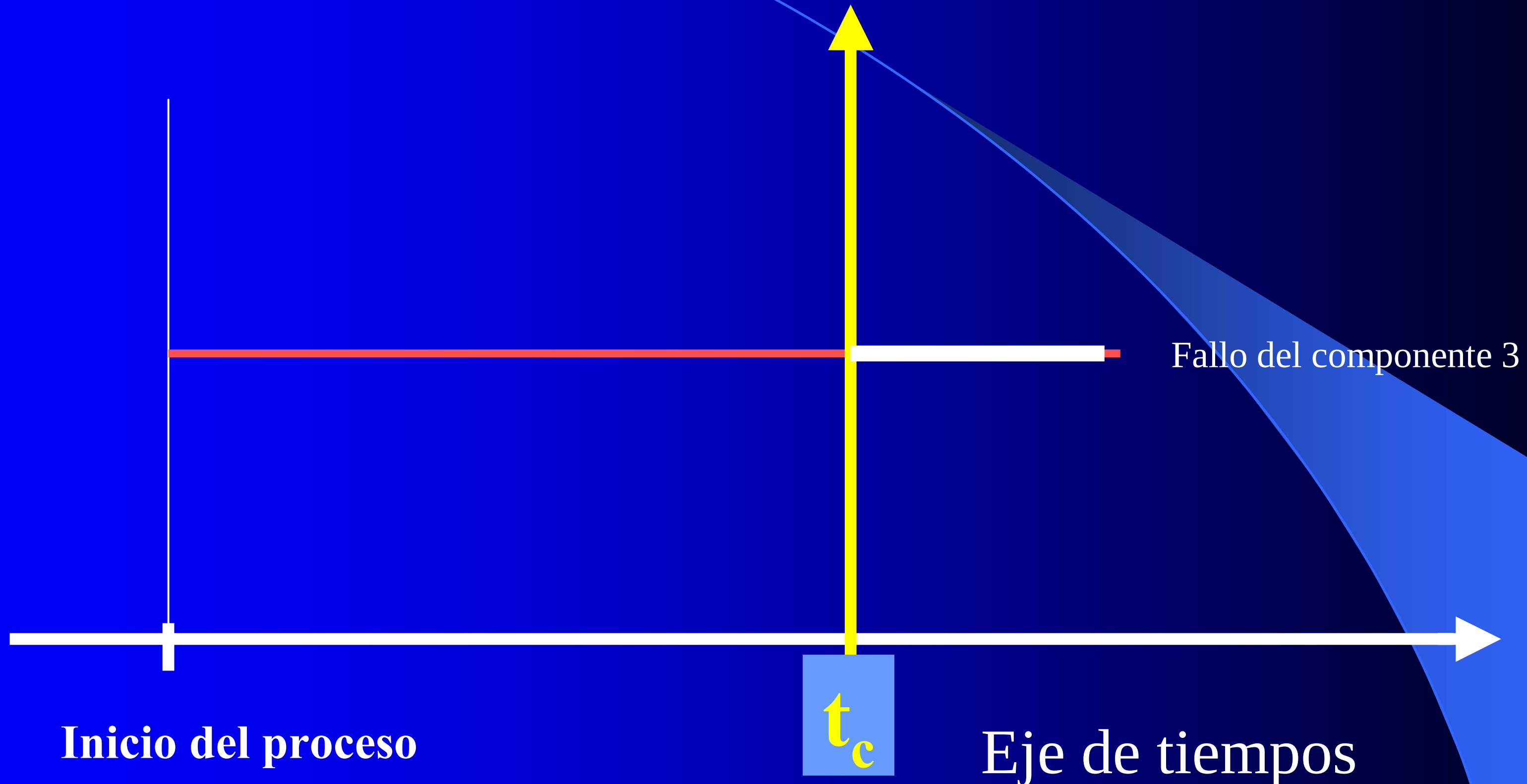
Fin del estudio al cabo de un tiempo t_c



Nuestros datos serán las duraciones de los tres primeros fallos: t_1, t_2 y t_4

PERO NO OBSERVAMOS EL FINAL DE t_3. Sólo sabemos que $t_3>t_c$

Fin del estudio al cabo de un tiempo t_c



Parte de la derecha no observada del componente 3
CENSURA POR LA DERECHA

Censura por la derecha

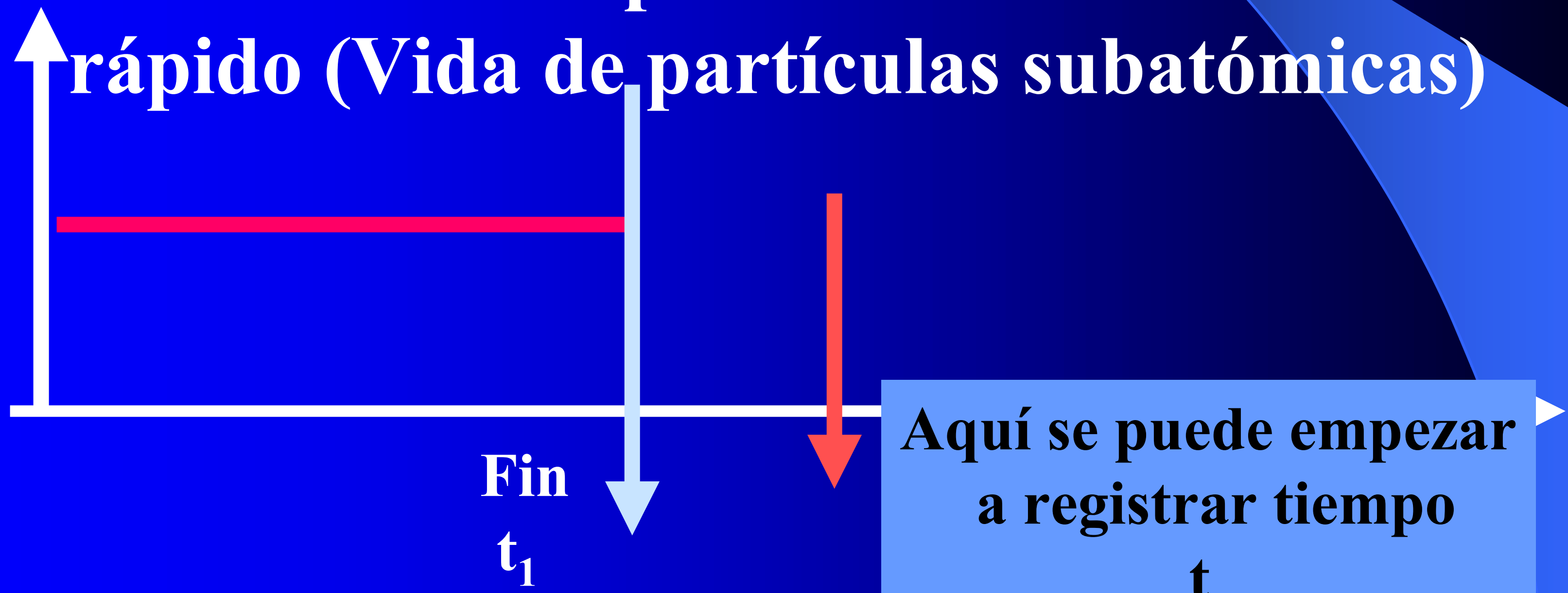
- **Economía:** duración del desempleo suele obtenerse de encuestas que preguntan a los parados cuanto tiempo llevan en paro. Al no conocerse el tiempo adicional que van a permanecer sin trabajo, solo se sabe su duración censurada.
  - El paro es superior al que el entrevistado indica en la encuesta. Si una persona dice que lleva en paro 3 meses, su paro real será $t_i > 3$

Censura por la derecha

- **Fiabilidad**: es muy normal poner a prueba una partida de componentes y observar los fallos durante un periodo de tiempo determinado. Los elementos que fallen durante este periodo proporcionaran observaciones completas. Los que sigan en funcionamiento al final del periodo proporcionaran observaciones censuradas.
  – El tiempo que se recoge para los elementos censurados será $t_i > t_c$

Censura por la izquierda

- **Cuando no podemos observar un acontecimiento por ocurrir demasiado rápido (Vida de partículas subatómicas)**



Registramos que la duración $t_1 < t_c$

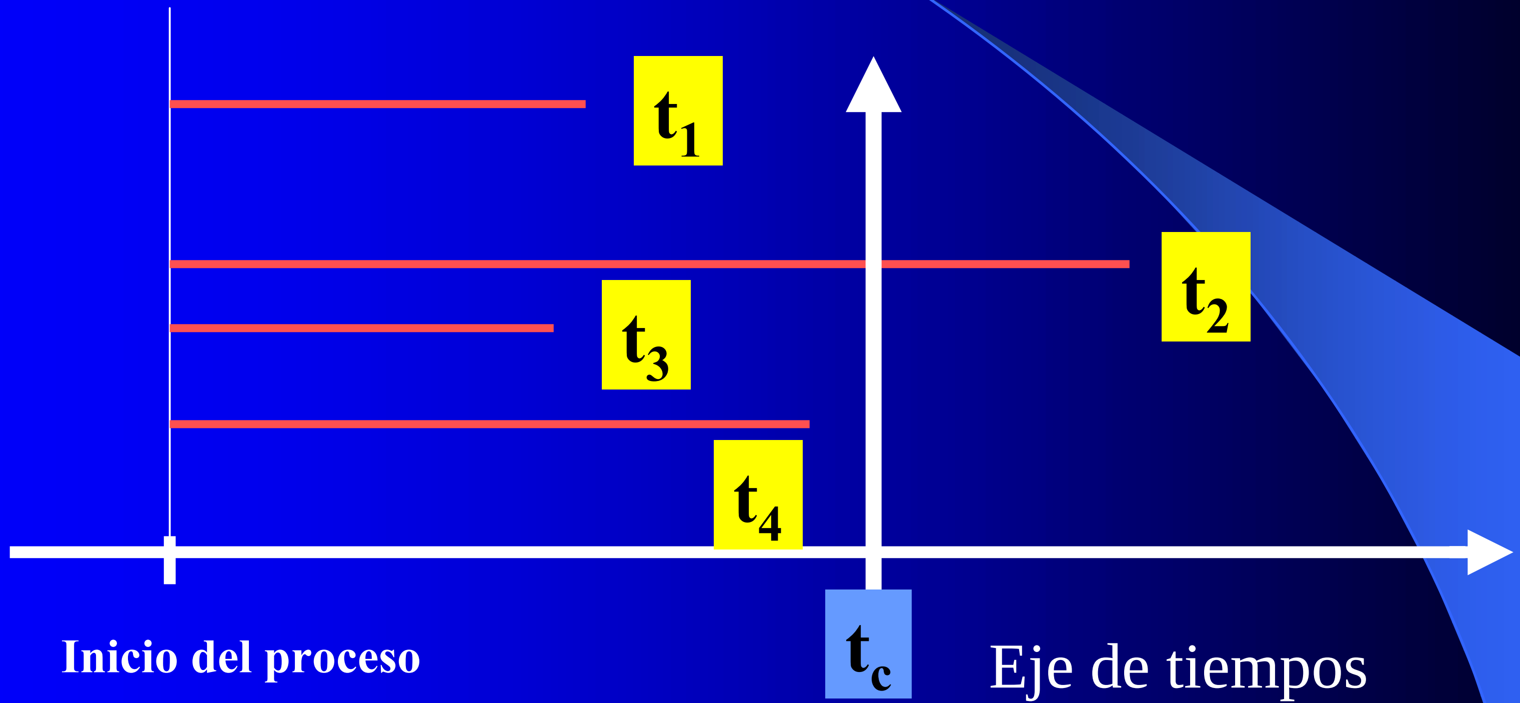
Censura por la izquierda

- Economía se producen censuras por la izquierda habitualmente. Un ejemplo son las edades de jubilación.
- Si tenemos como dato la edad de una persona y sabemos que esta jubilada, podemos deducir que su edad de jubilación es menor que su edad actual
  - Registramos Ejub<Edad

Tipos de censura

- **Tipo 1:**
  - **El experimento que genera datos con censura de tipo 1 consiste en poner a prueba una partida de n componentes y observarlos durante un tiempo predeterminado t_c**
  - **La duración t_c es decidida por el experimentador.**
  - **Se observan los datos completos correspondientes a los r componentes que han fallado antes de t_c.**
  - **La duración de los n-r componentes que no han fallado sabemos que es mayor que t_c**

Censura de Tipo 1

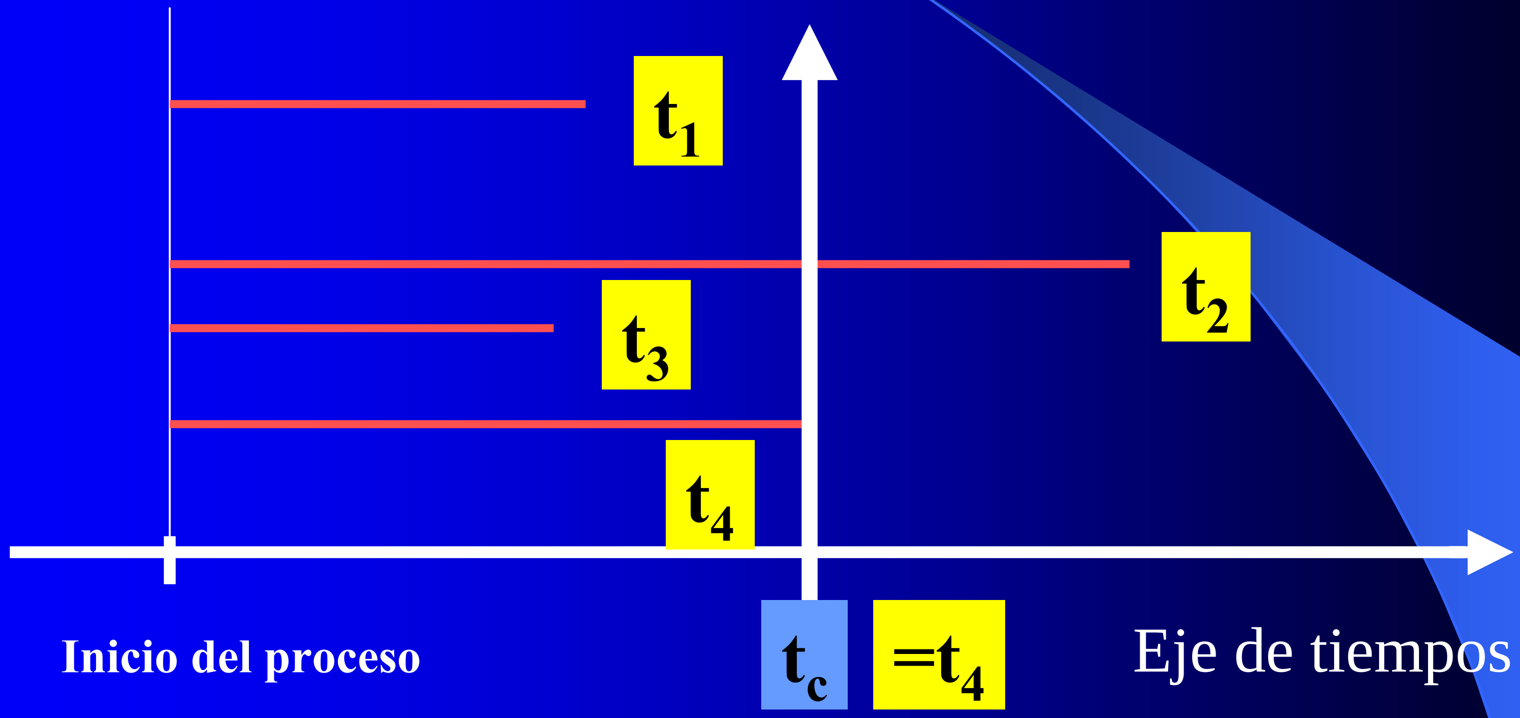


Nuestros datos serán las duraciones de estos cuatro componentes:
t_1, t_3 y t_4 Además $t_2 > t_c$.
n=4 y r=3

Tipo 2

- **Tipo 2:**
  - El experimento que genera datos con censura de tipo 2 consiste en poner a prueba una partida de n componentes y observarlos hasta que falla el r-ésimo componente en el instante t_c
  - El número de fallos es decidido previamente por el experimentador.
  - La duración t_c NO es decidida por el experimentador.
  - Se observan los datos completos correspondientes a los r componentes que han fallado antes de t_c.
  - La duración de los n-r componentes que no han fallado sabemos que es mayor que t_c

Censura de Tipo 2.
Termina cuando falle el 75%

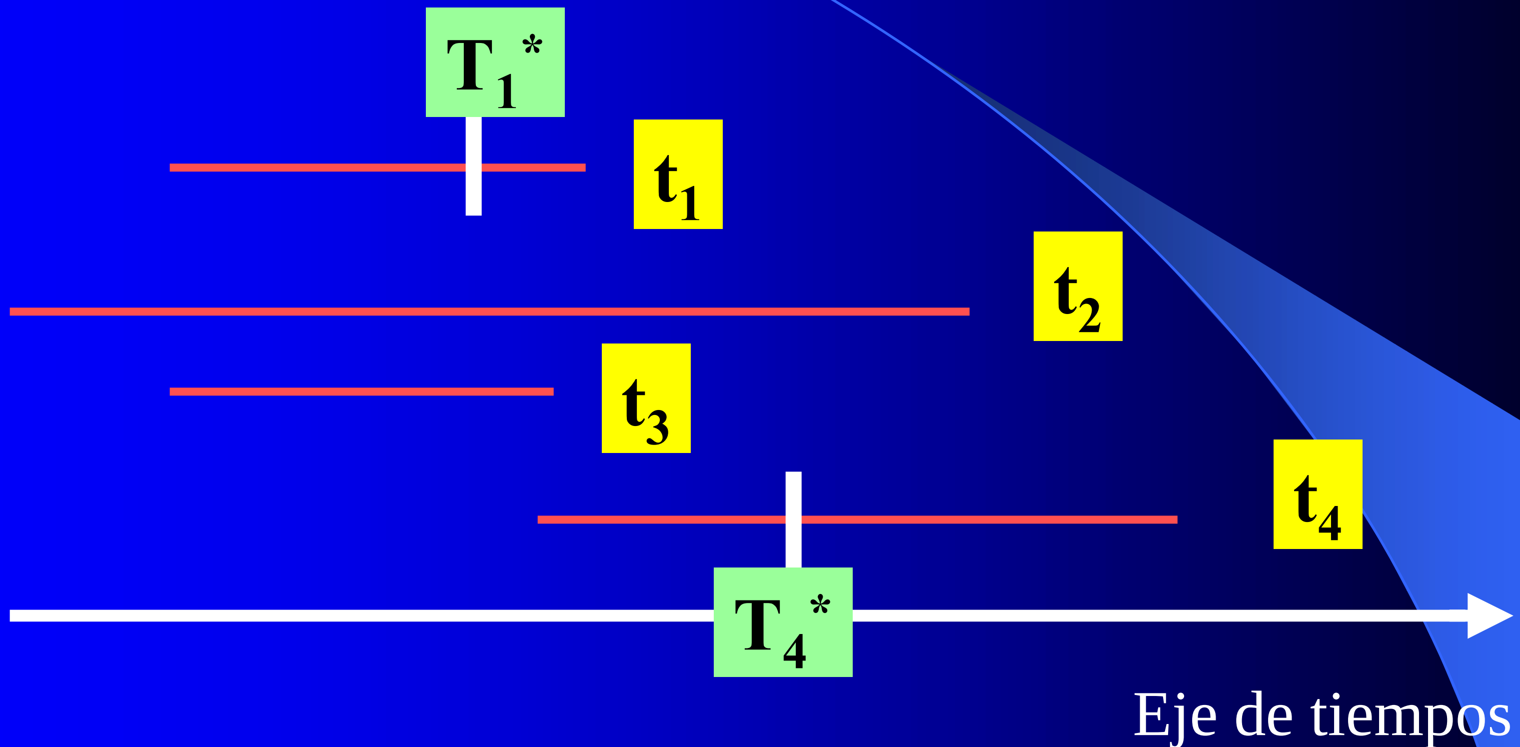


Nuestros datos serán las duraciones de estos cuatro componentes:
t_1, t_3 y t_4 Además $t_2 > t_4$.
n=4 y r=3

Censura aleatoria

- La censura se produce aleatoriamente
- Se trata a un grupo de pacientes con un nuevo tratamiento que mejora su supervivencia a determinada enfermedad. Un paciente se traslada de ciudad y no vuelve al control del hospital. Veremos una serie de observaciones completas y otras censuradas.

Censura aleatoria



Nuestros datos serán las duraciones de estos cuatro componentes:

t_2 y t_3 completos

Además t_1 y t_4. Fallan o desaparecen antes del final de su tiempo.

Observamos $T_1^*<t_1$ y $T_4^*<t_4$ $\quad$ n=4 y r=2

Estimación con datos censurados

- **Mucho más compleja que con datos completos**
- **En general es imposible calcularlo a mano**
- **Lo haremos en ordenador**

Estimación con datos censurados

- Si hay censura las técnicas descriptivas básicas no van a servir.
- No podremos realizar histogramas si no conocemos la longitud final de las observaciones.
- La única forma de saber qué modelo elegir es usar los gráficos a escala que hemos aprendido pero adaptados a la censura
- El gráfico usado es el estimador de producto límite o estimador de Kaplan Meier

Estimador de Kaplan Meier

- Estima la función de supervivencia cuando hay censura

- Vamos a estudiarlo con un ejemplo

Estimador de Kaplan Meier

- Se realiza un experimento para saber si una nueva droga es efectiva tratando una enfermedad mortal.
- Un grupo de pacientes es tratado con la nueva droga (6M) y el otro con placebo.
- El ensayo es doble ciego
- Datos de primer grupo (6MP):
  - 6 6 6 6* 7* 9 10* 10* 11 13 16* 17* 19* 20 22 23* 25* 32* 32* 34* 35
- Datos del segundo grupo (Placebo)
  - 1 1 2 2 3 4 4 5 5 8 8 8 8 11 11 12 12 15 17 22 23

Estimador de Kaplan Meier

- **Realiza las siguientes operaciones:**
  - **Se ordenan los valores de menor a mayor**
  - **Para cada tiempo de fallo (Si hay varios fallos en el mismo momento, para el ultimo) se calcula el numero de individuos que quedan en riesgo.**
  - **El estimador para el primer tiempo de fallos será:**
    - $S(t_1)=(n_1-d_1)/n_1$
  - **n_1 representa el numero de individuos que están en riesgo justo antes del primer tiempo de fallo.**
  - **d_1 es el número de fallos/muertes en el primer tiempo de fallo.**
  - **Para el segundo tiempo de fallo será**
    - $S(t_2)=[(n_2-d_2)/n_2].S(t_1)$
    - $S(t_3)=[(n_3-d_3)/n_3].S(t_2)$

6MP): 6 6 6 6* 7* 9 10* 10* 11 13 16* 17* 19* 20 22 23* 25* 32* 32* 34* 35

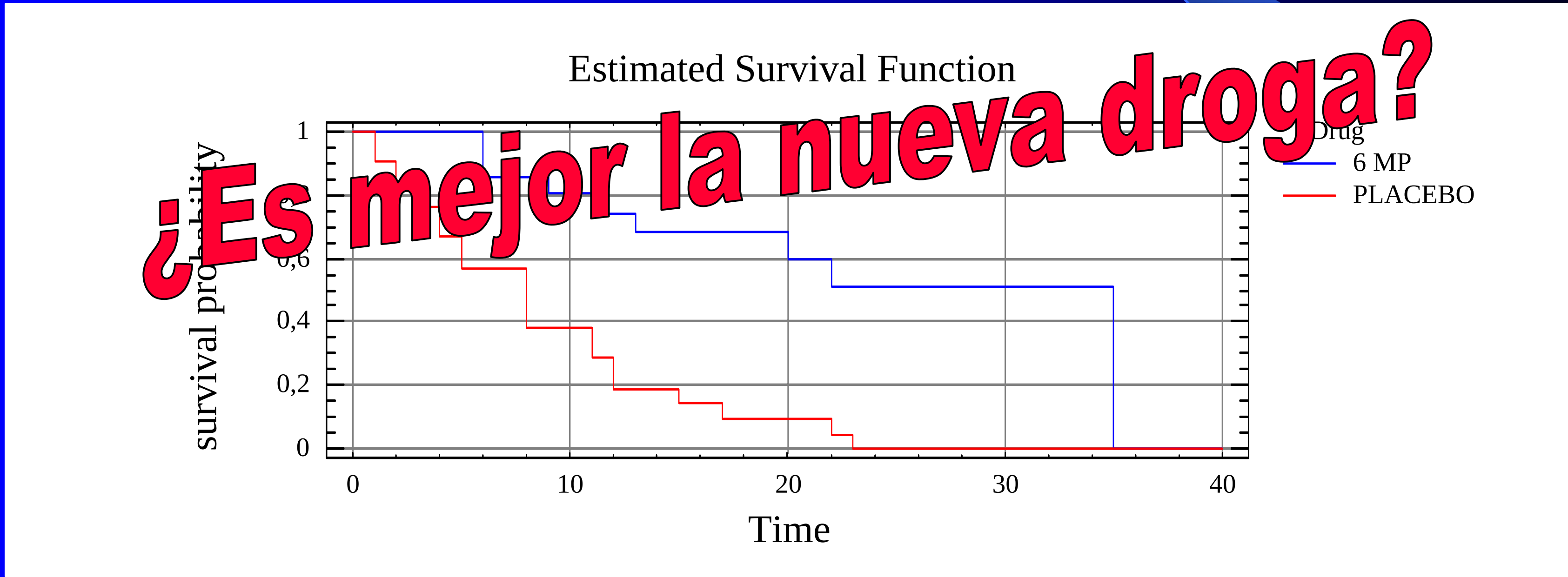
Sólo se registran Tiempos de FALLO
No censuras. Por eso no aparece 7 o 9

Tiempo	Riesgo	Fallos	$\frac{n_i-d_i}{n_i}$	S(t)
6	21	3	$\frac{21-3}{21}$	$\frac{21-3}{21} = 0.8571$
9	16	1	$\frac{16-1}{16}$	$\frac{16-1}{16}0.8571 = 0.8036$
11	13	1	$\frac{13-1}{13}$	$\frac{13-1}{13}.0.8036 = 0.7418$
13	12	1	$\frac{12-1}{12}$	$\frac{12-1}{12}0.7418 = 0.6799$
20	8	1	$\frac{8-1}{8}$	$\frac{8-1}{8}.0.6799 = 0.5950$
22	7	1	$\frac{7-1}{7}$	$\frac{7-1}{7}.0.5950 = 0.51$
35	0	1	0	0

Ordenador

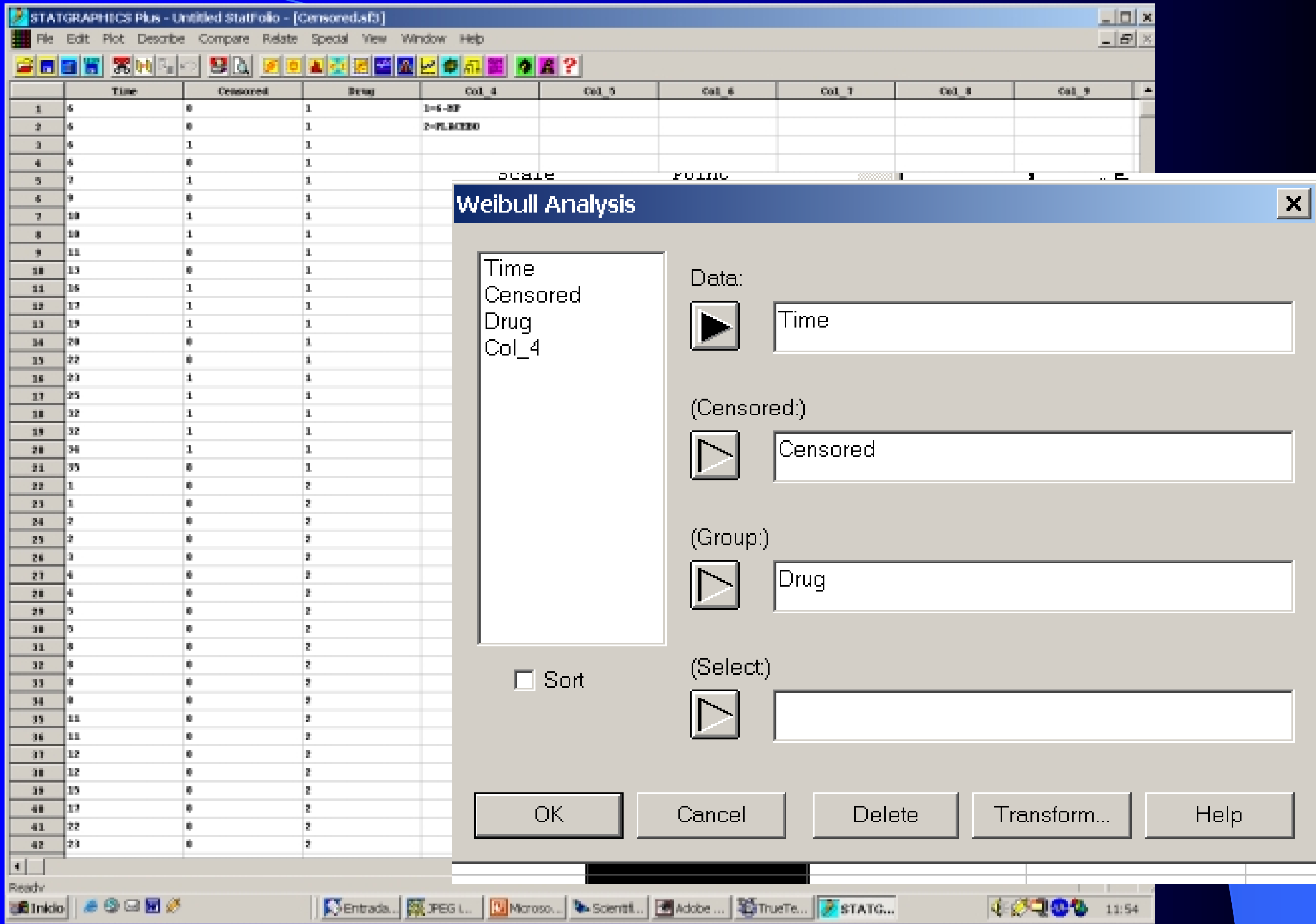
- Statgraphics hace este estimador en
  - DESCRIBE
  - DISTRIBUTION FITTING Y LIFE TABLES (TIMES)
  - Se escriben los tiempos y se añade una variable de censura que toma el valor 0 si la variable es completa y el valor 1 si es censurada

Ordenador


Estimated Survival Function
¿Es mejor la nueva droga?
survival probability
1
0,6
0,4
0,2
0
0
10
20
30
40
Time
Drug
6 MP
PLACEBO

Estimación paramétrica con censura (Análisis Weibull)

- El proceso con STATGRAPHICS. es el siguiente:
  - DESCRIBE
  - Distribution Fitting (censored data) y Weibull Analysis
  - Se escriben los tiempos y se añade una variable de censura que toma el valor 0 si la variable es completa y el valor 1 si es censurada
  - Se añade una variable de grupo si lo hay


STATGRAPHICS Plus - Untitled StatFolio - [Censored.sf3]
File Edit Plot Describe Compare Relate Special View Window Help
Weibull Analysis
Time
Censored
Drug
Col_4
Sort
Data:
Time
(Censored:)
Censored
(Group:)
Drug
(Select:)
OK
Cancel
Delete
Transform...
Help

```
Analysis Summary

Data variable: Time
Censoring: Censored
Groups: Drug
Number of groups = 2
Estimation method: maximum likelihood
```

Group	Sample Size	Number of Failures	Estimated Shape	Estimated Scale	Startin Point
1	21	9	1,50716	32,3886	0,0
2	21	21	1,3705	9,48214	0,0

$$\widehat{S(t)} = \mathrm{exp}(-(32.8t)^{1.5})$$

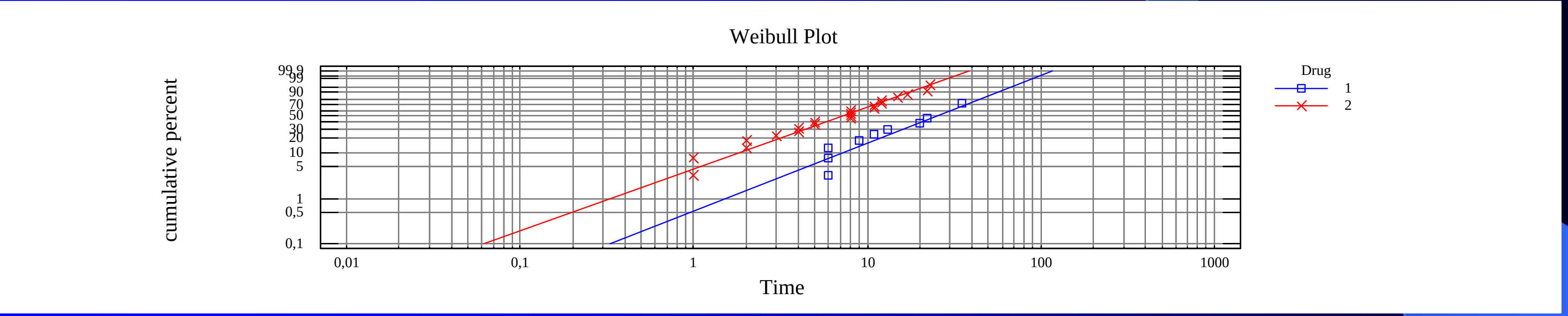
para la droga 6MP

$$\widehat{S(t)} = \mathrm{exp}(-(9.48t)^{1.37})$$

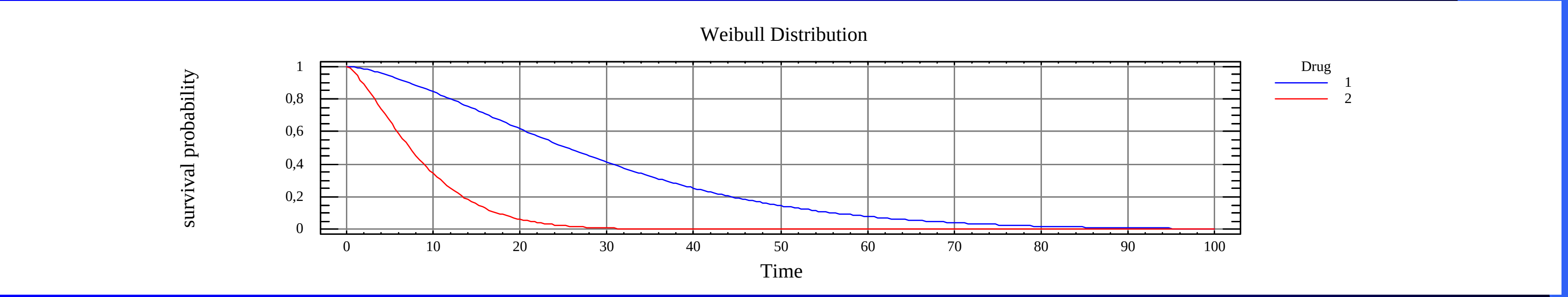
para el placebo.

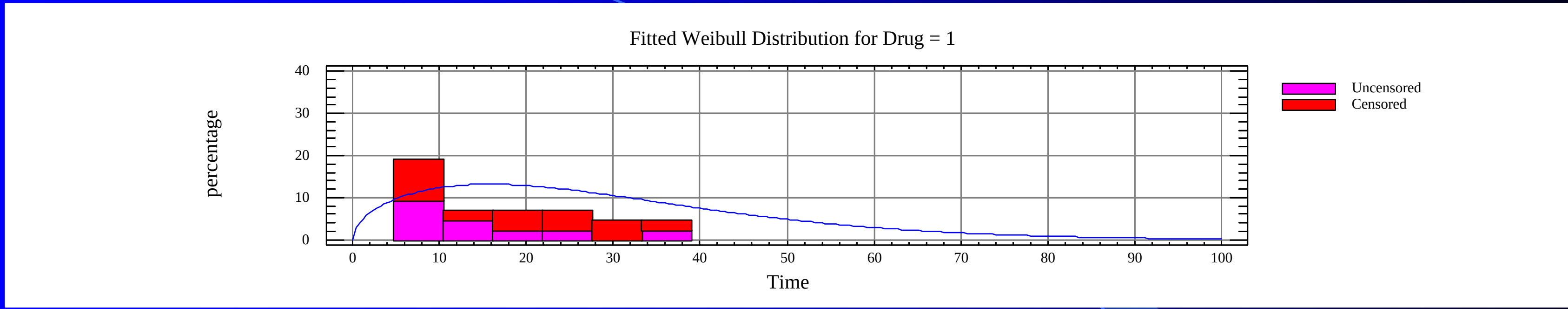
```
Analysis Summary

Data variable: Time
Censoring: Censored
Groups: Drug
Number of groups = 2
Estimation method: maximum likelihood
```

Group	Sample Size	Number of Failures	Estimated Shape	Estimated Scale	Starting Point
1	21	9	1,50716	32,3886	0,0
2	21	21	1,3705	9,48214	0,0

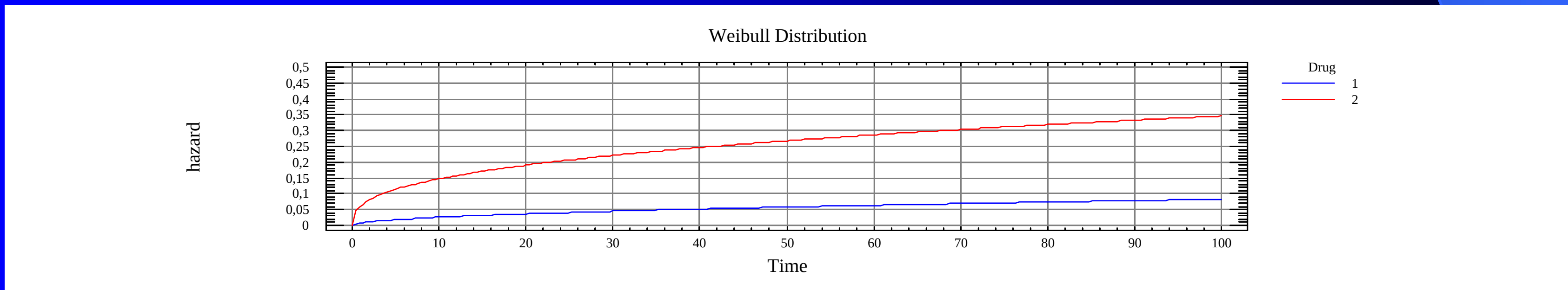


Están alineados así que son Weibull

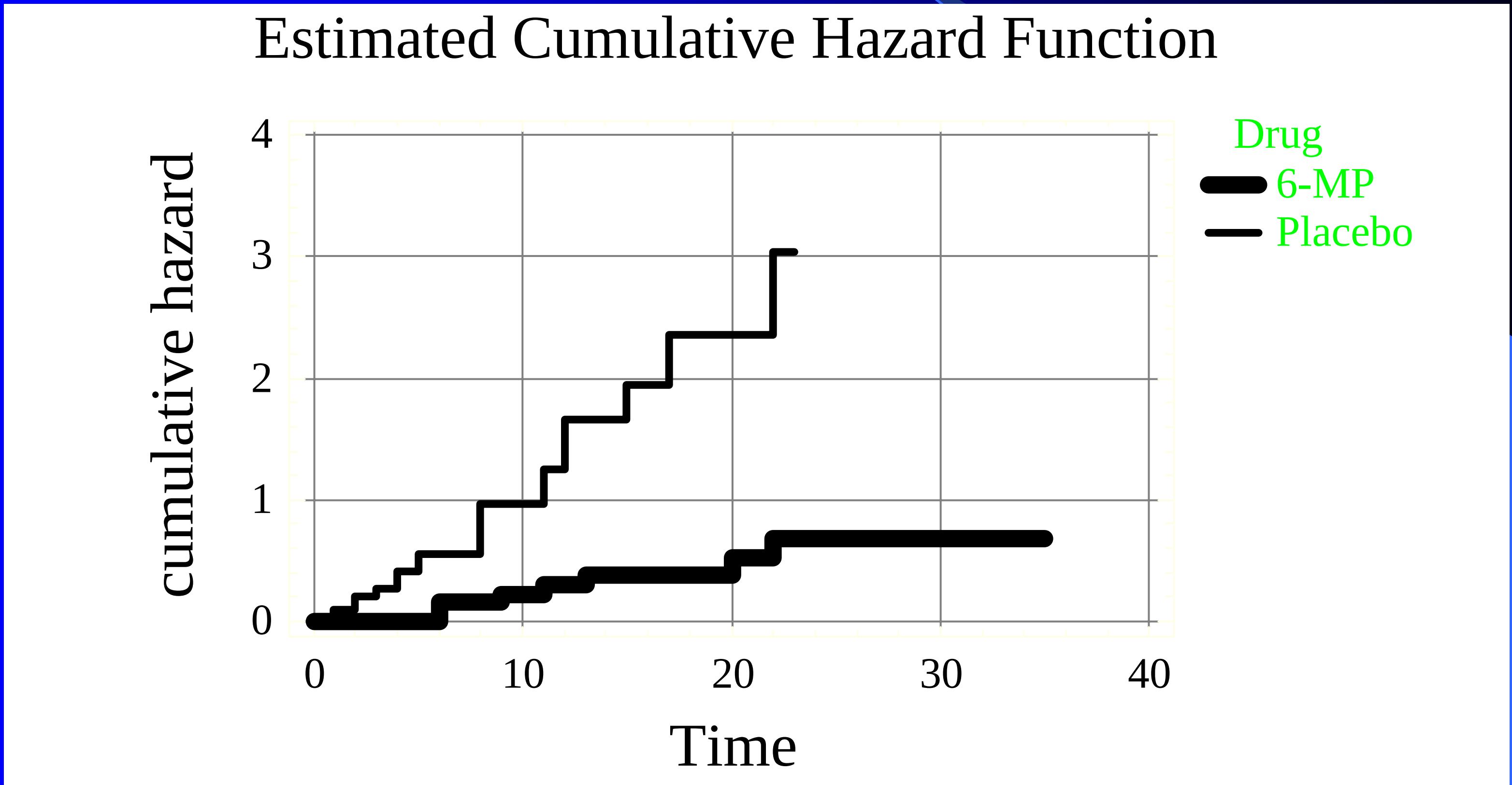




Parece que no ajusta: Es por la censura que no la tiene en cuenta
En el papel Weibull estaban alineados



Ordenador


Estimated Cumulative Hazard Function
cumulative hazard
0
1
2
3
4
0
10
20
30
40
Time
Drug
6-MP
Placebo

Ensayos acelerados

Ensayos acelerados

- Surgen debido a que algunos productos tienen unas duraciones tan elevadas que es imposible seguir un experimento hasta el final.
- Por ejemplo componentes diseñados para durar 40 años. Es muy improbable que alguno falle en el tiempo en que razonablemente se puede realizar un ensayo.
- Se pone a prueba el componente bajo condiciones de trabajo mucho mas desfavorables de las habituales y se propicia que el fallo se produzca antes

Ensayos acelerados

- La realización de ensayos acelerados es compleja y debe ser planificada por los propios ingenieros de diseño, ya que hay que tener en cuenta que factores hay que acelerar y en que medida.
- Por ejemplo, si queremos acelerar un ensayo con válvulas de precisión, Será preciso determinar si acelerar la presión de trabajo, la temperatura o la concentración de elementos oxidantes.

El esquema de trabajo es el siguiente:

- **Se obtienen datos de tiempos de fallo con diversas aceleraciones.**
- **Se estima mediante un análisis Weibull la distribución para cada uno de esos niveles**
- **Se calcula la mediana y los percentiles 10% y 90%.**
- **Se dibuja en un grafico la mediana y los percentiles respecto al nivel de stress**
- *Se extrapola para las condiciones nominales.*

Se extrapola para las condiciones nominales.

Si no hay experiencias previas
la extrapolación es siempre peligrosa

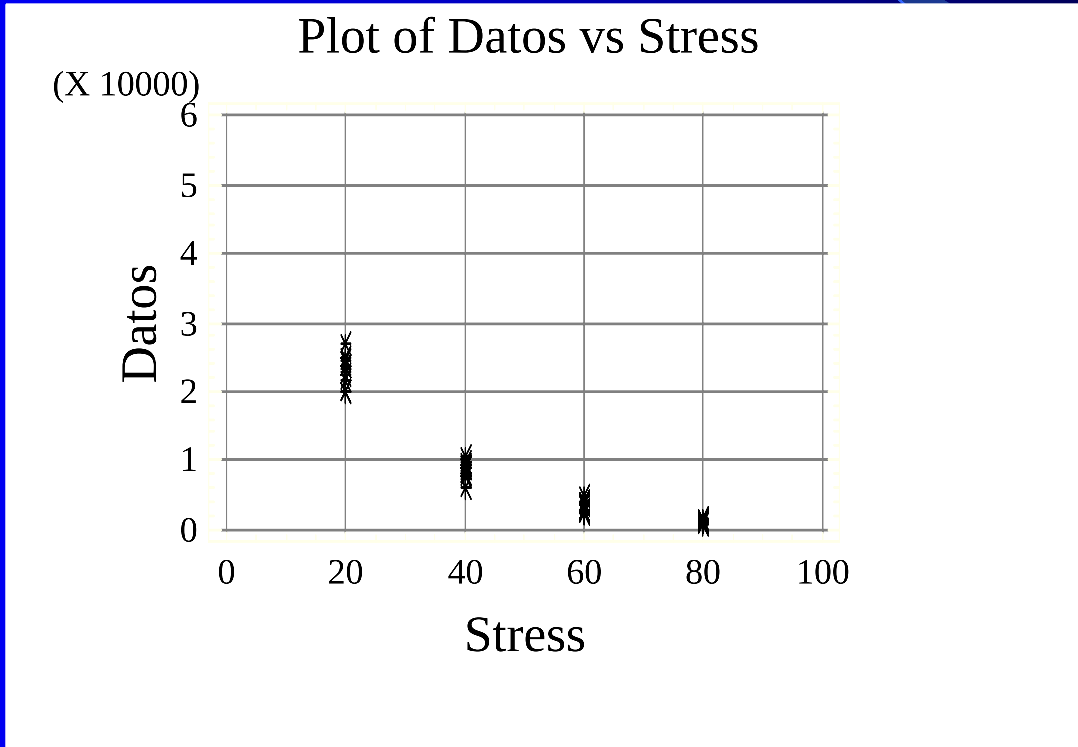
Ejemplo

- Los datos representan tiempos de fallo en horas de un componente en función de su stress. El componente debe funcionar en condiciones de Stress=4. Los datos con asterisco son censurados.

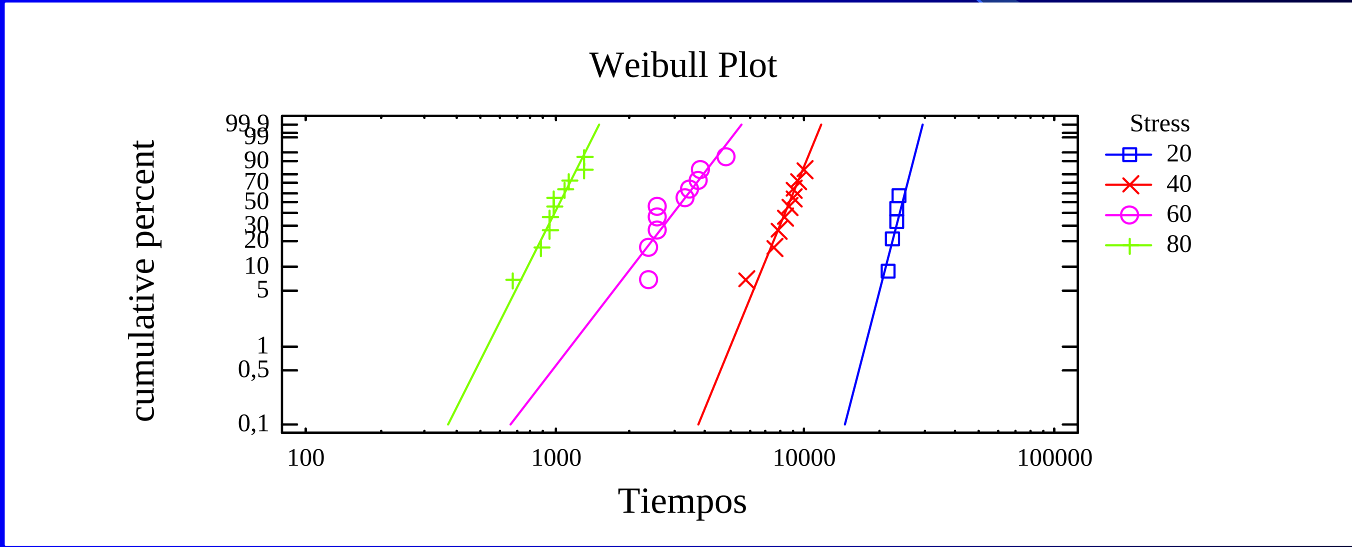
Datos

Stress=20	Stress=40	Stress=60	Stress=80
23567	8674	3456	1300
21569	5780	2567	879
24852*	9007	2374	679
22500	10345*	2567	1300
26753*	8909	3346	956
24456*	7890	2568	980
23789	9870	3789	1098
19876*	7479	3768	1134
19889*	9345	4896	956
23456	8234	2365	984

Representación gráfica: Datos en función del Stress. Hay datos censurados.



Análisis Weibull



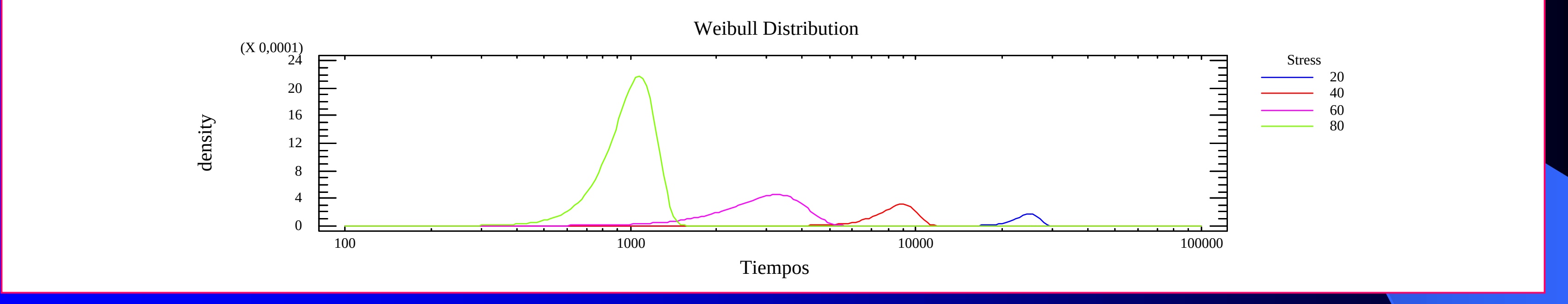
Datos alineados en los cuatro grupos

Análisis Weibull

Group	Sample Size	Number of Failures	Estimated Shape	Estimated Scale	Starting Point
20	10	5	12,4232	25379,7	0,0
40	10	9	7,68861	9137,44	0,0
60	10	10	4,15438	3481,62	0,0
80	10	10	6,42138	1101,88	0,0

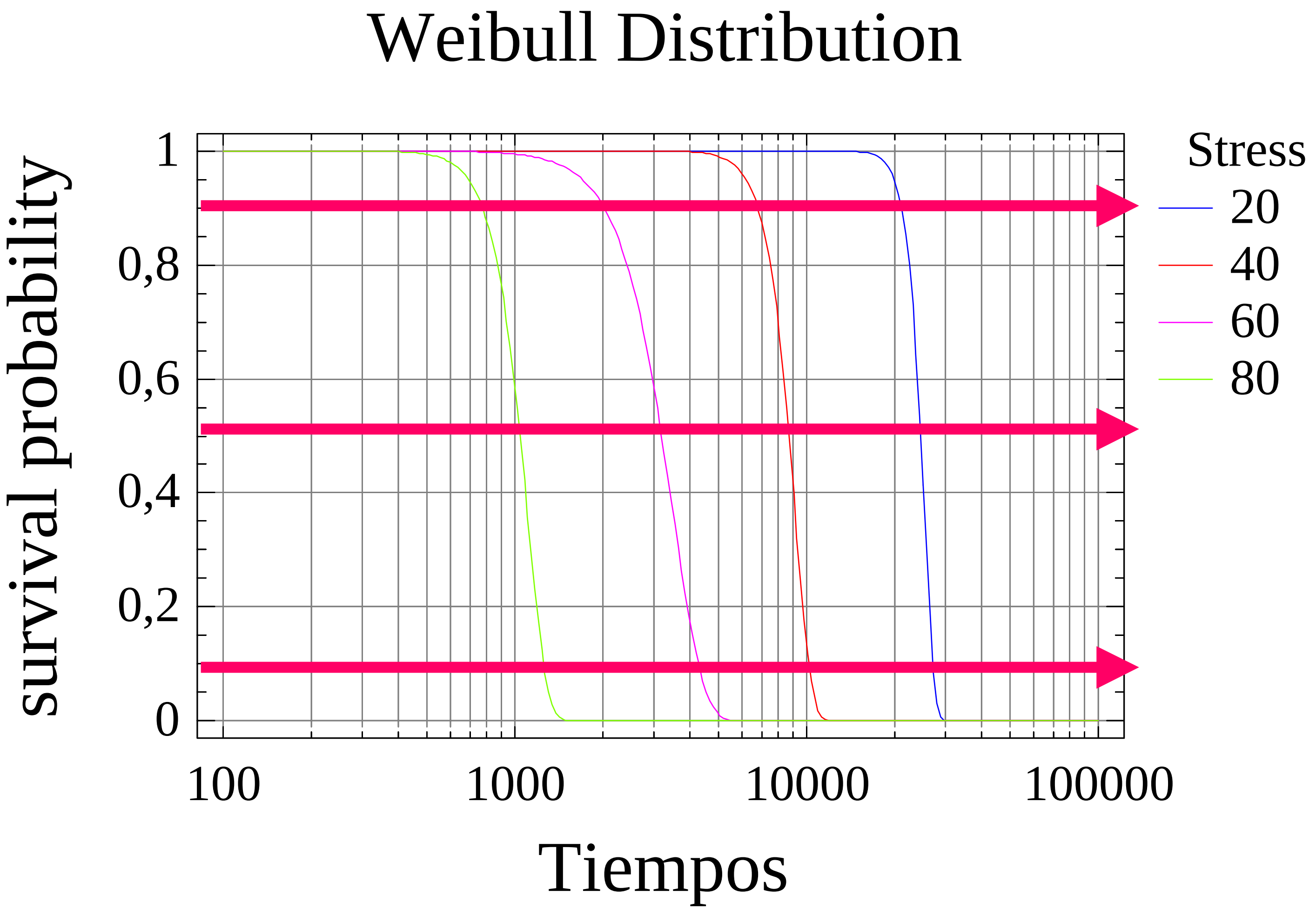
Valores de lambda y Beta para los cuatro grupos

Análisis Weibull



Las cuatro funciones de densidad

Supervivencias



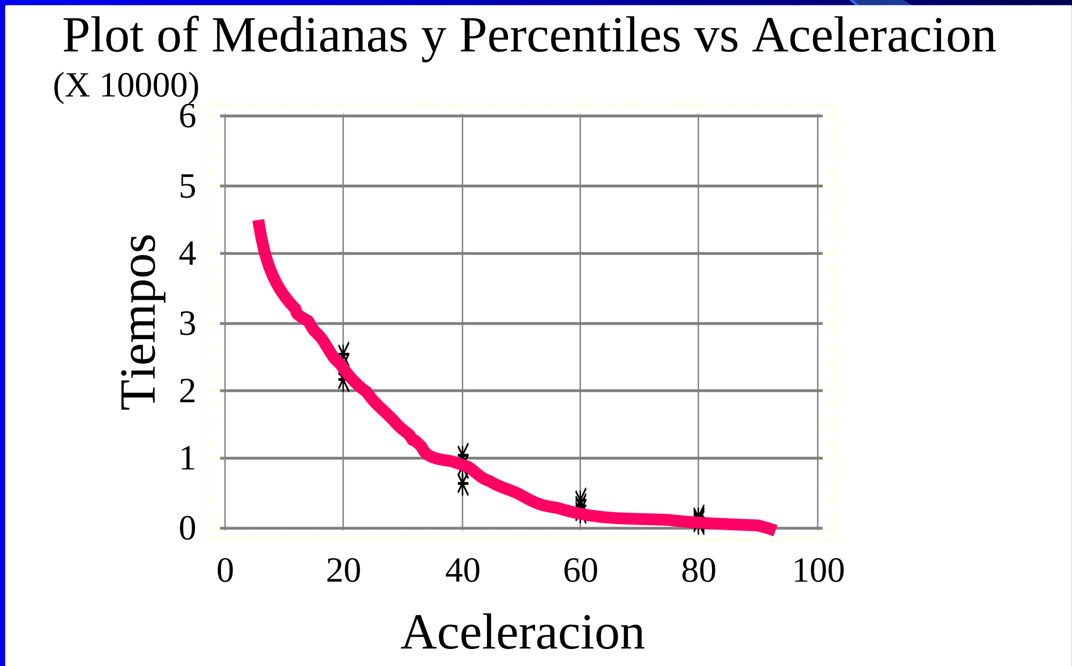
Marcamos los percentiles 10, 50 y 90

Percentiles (critical values)

Critical Values for Tiempos

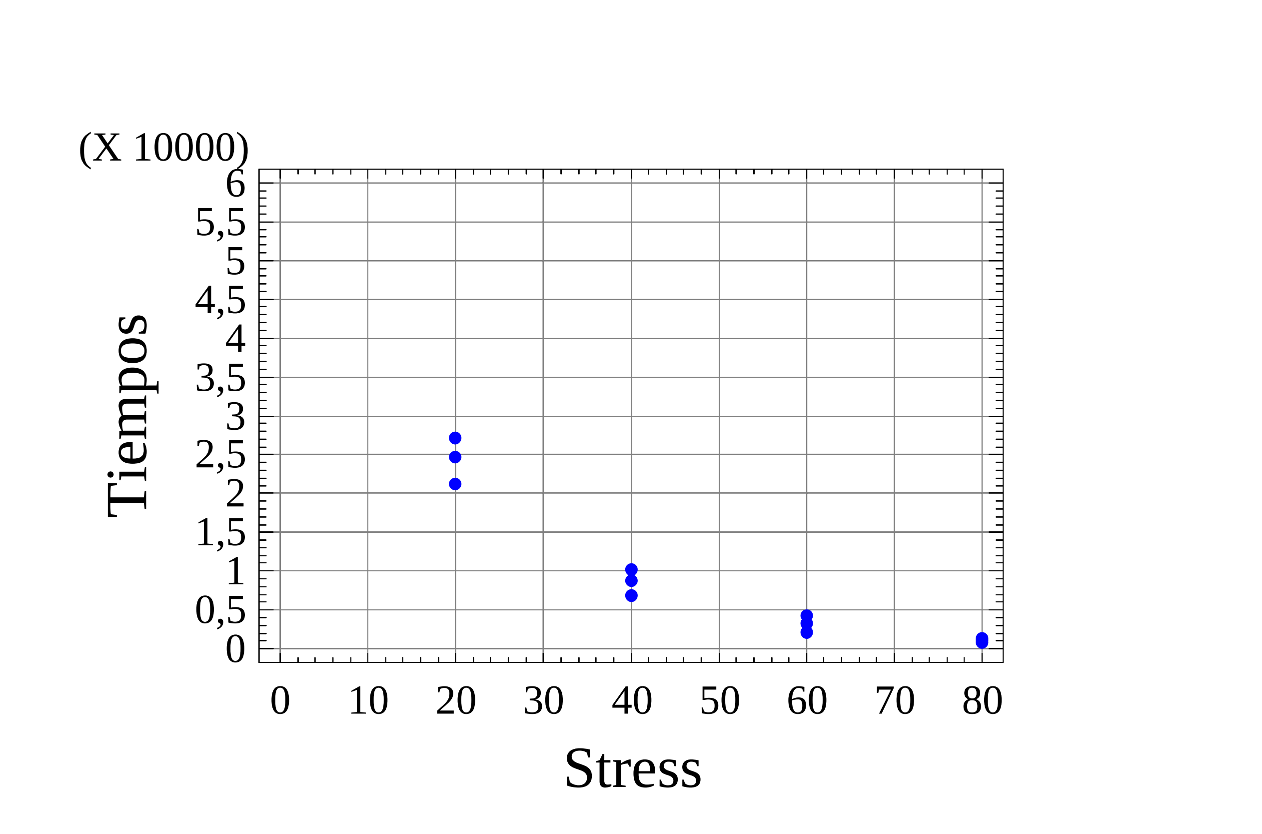
Group	Lower Tail Area	Critical Value
20	0,1	21174,7
	0,5	24641,8
	0,9	27142,0
40	0,1	6818,85
	0,5	8712,08
	0,9	10184,4
60	0,1	2025,49
	0,5	3187,62
	0,9	4255,69
80	0,1	776,135
	0,5	1040,75
	0,9	1254,71

Haciendo un gráfico

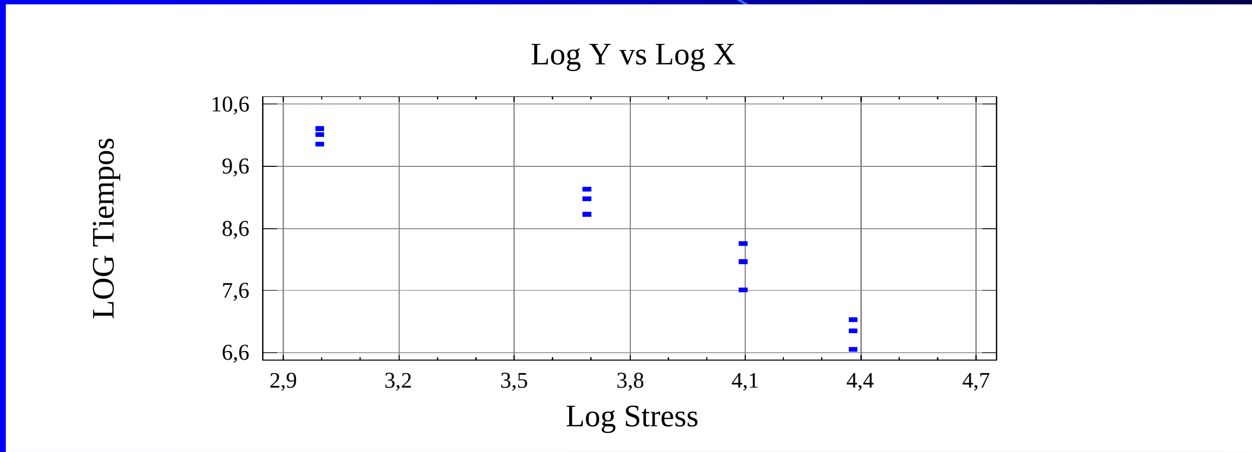


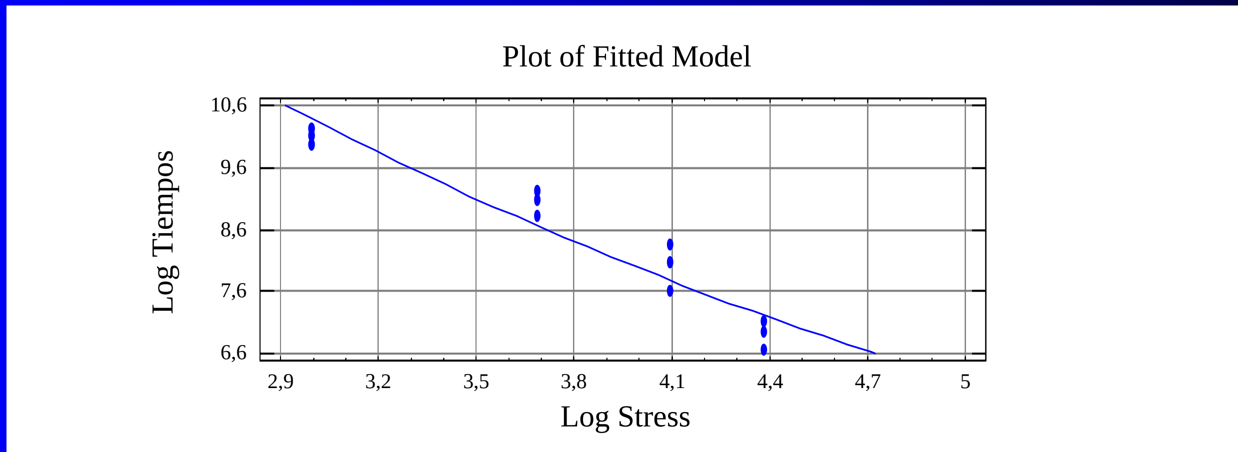
Vemos que ajustar una exponencial sería adecuado

Regresión exponencial: plot de los datos

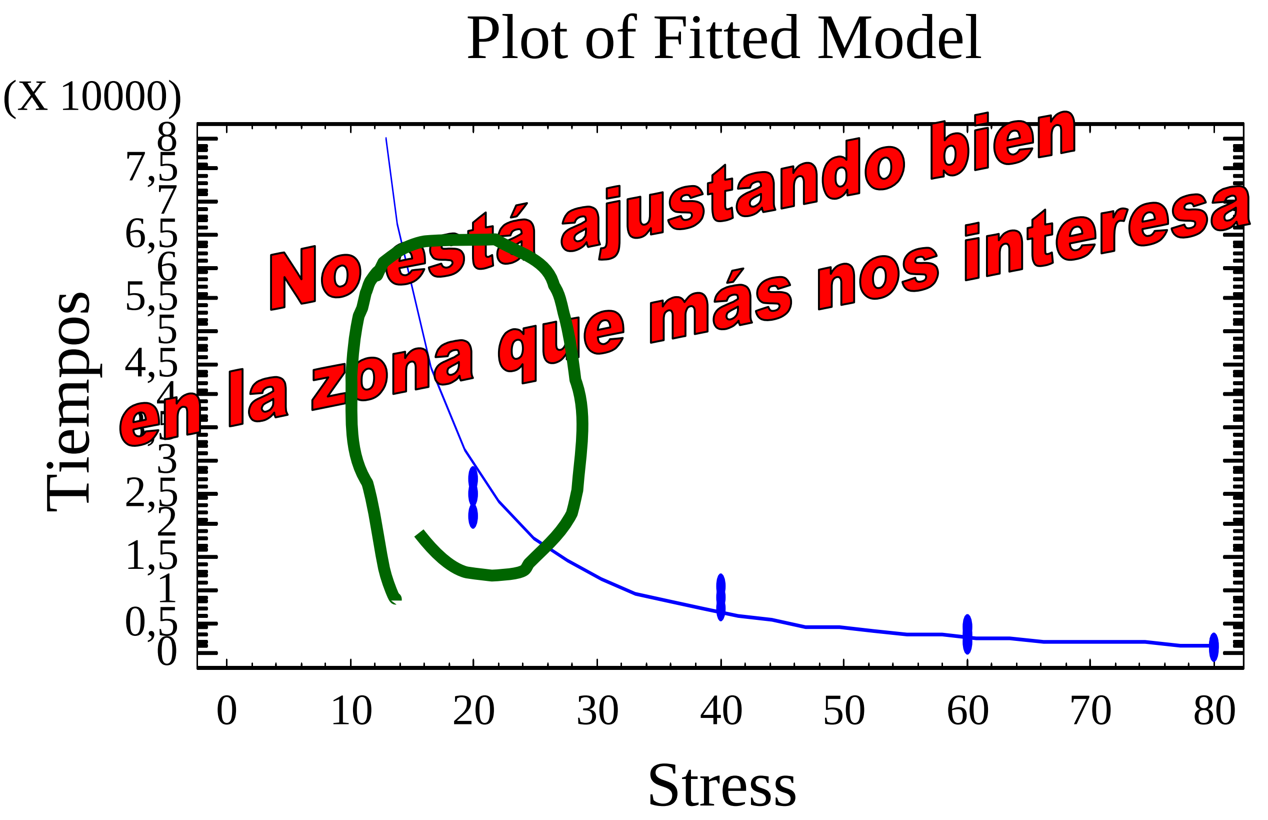


Transformando ambas variables a logaritmos

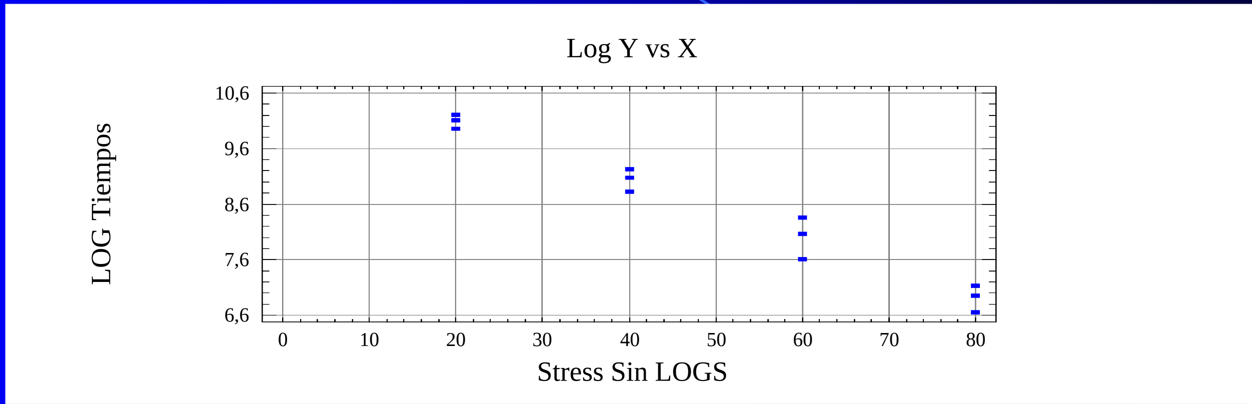


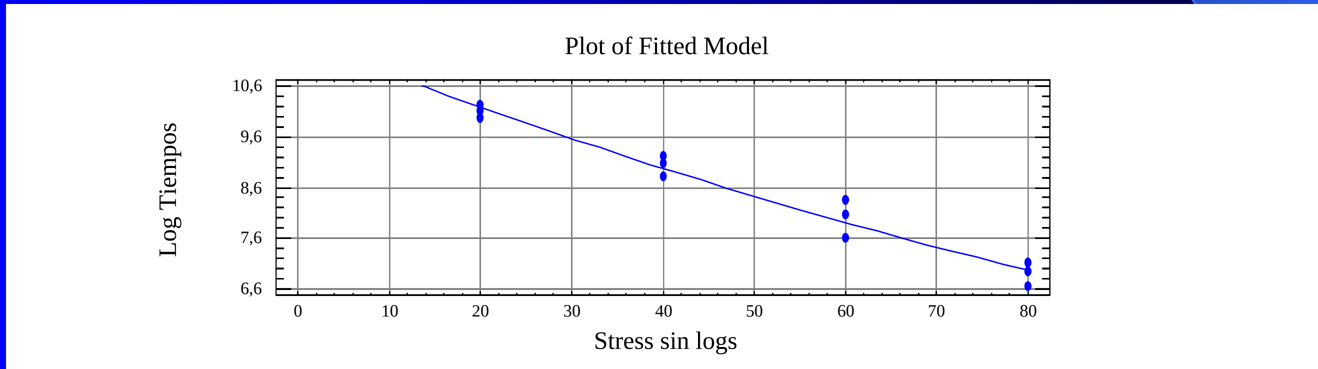


Quitando los logs:
tiempo=beta1.stress^{beta2}

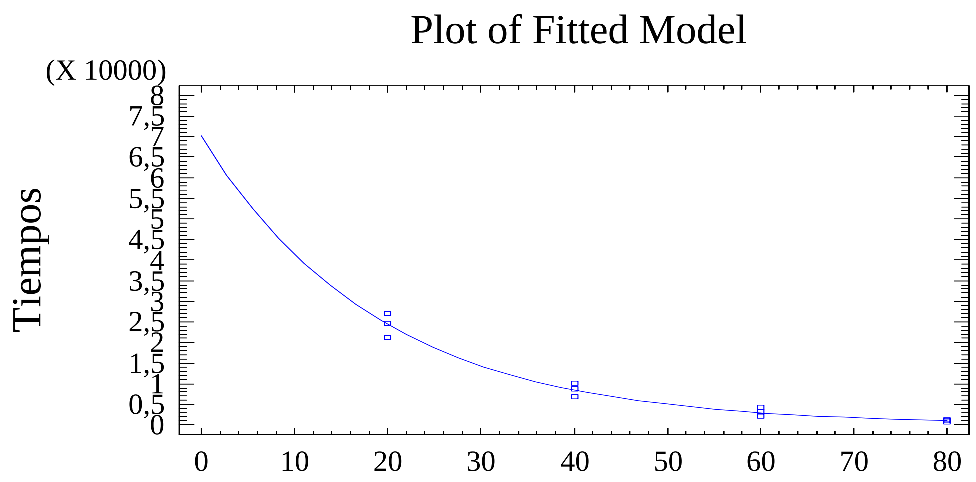


Transformando sólo Y a logaritmos





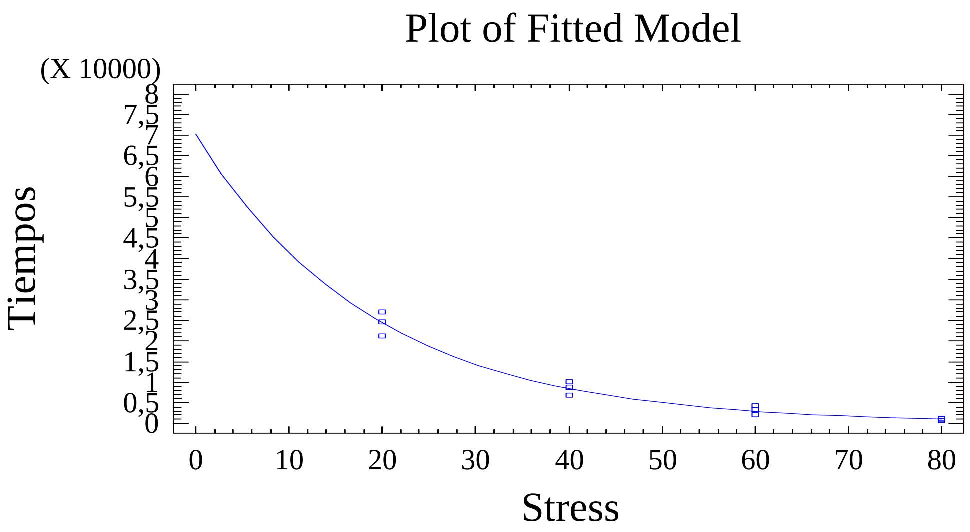
Mucho más lineal



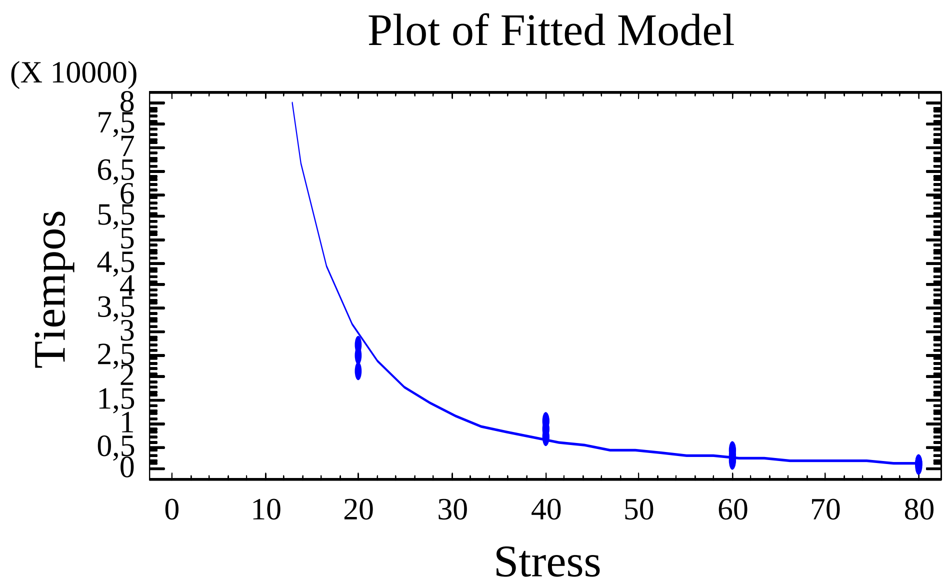

```
Regression Analysis - Exponential model: Y = exp(a + b*X)
-----------------------------------------------------------------------------
Dependent variable: Col_5
Independent variable: Col_4
-----------------------------------------------------------------------------
                               Standard          T
Parameter       Estimate         Error       Statistic        P-Value
-----------------------------------------------------------------------------
Intercept        11,1592       0,160231        69,6442         0,0000
Slope         -0,0528778     0,00292541       -18,0753         0,0000
-----------------------------------------------------------------------------
```

$$\text{Tiempos} = 11.16\ \text{stress}^{-0.05}$$

Modelo con un log en Y



Modelo con dos logs



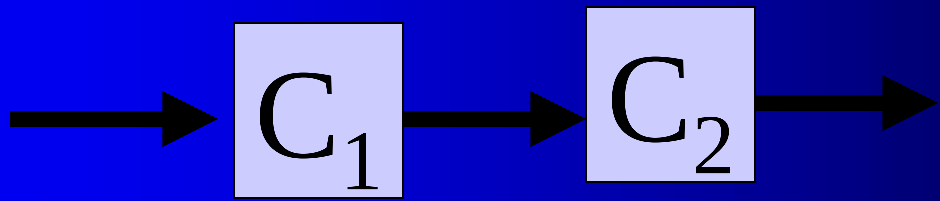
Ojo a las extrapolaciones

Fiabilidad de Sistemas

Sistemas

- **Hemos estudiado cómo estimar la Fiabilidad/Duración de componentes simples**
- **En la práctica están integrados en sistemas más complejos.**
- **Dentro de sistemas complejos los más elementales son los sistemas serie y paralelo**

Sistema Serie



El sistema no funciona cuando el flujo de señal entre la entrada y la salida se interrumpe

Es decir sólo funciona si funcionan los dos componentes

$S_1(t)$ **Función de supervivencia del Componente 1**

$S_2(t)$ **Función de supervivencia del Componente 2**

$$P_s\,(Funcione) = P_1(Funcione)xP_2(Funcione) =$$

$$S_s(t) = S_1(t).S_2(t)$$

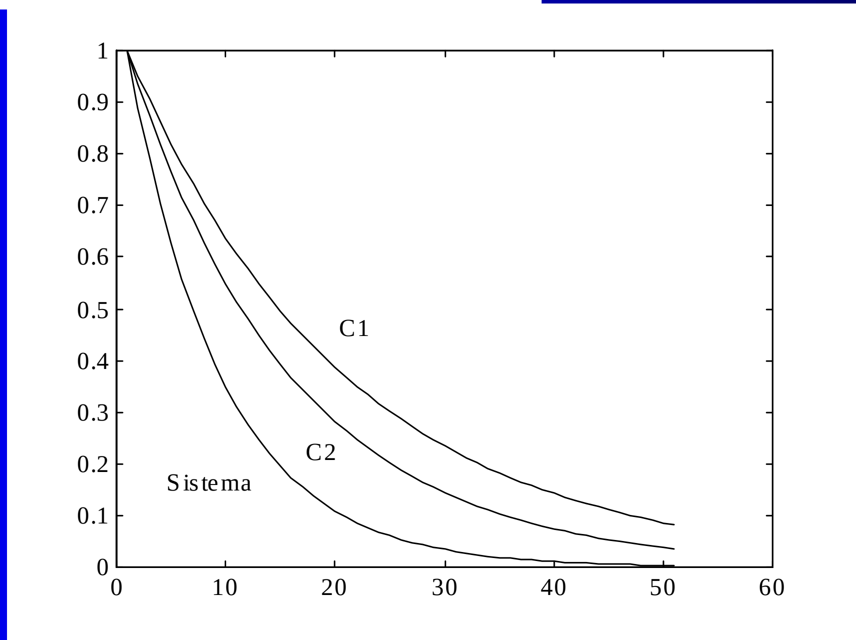
Ejemplo

- Un sistema serie con dos componentes.
- $S_1(t)=\exp(-t/2000)$
- $S_2(t)=\exp(-t/1500)$
- Vamos a Calcular la fiabilidad del sistema serie.
- $S_S(t)= S_1(t\ ).\ S_2(t\ )=\exp(-t/2000).\ \exp(-t/1500)=$
- $S_S(t)=\exp(-t/200-t/1500)=\exp(-t/857)$

$$S_S(t)=\exp(-t/857)$$

LA FIABILIDAD DEL SISTEMA SERIE ES MENOR QUE LA DE CUALQUIERA DE SUS COMPONENTES.

$S_S(t)=\exp(-t/857)$



Las probabilidades de que los componentes duren más de1000 horas son

$$S_1(t) = \exp(-1000/2000) = 0.61$$

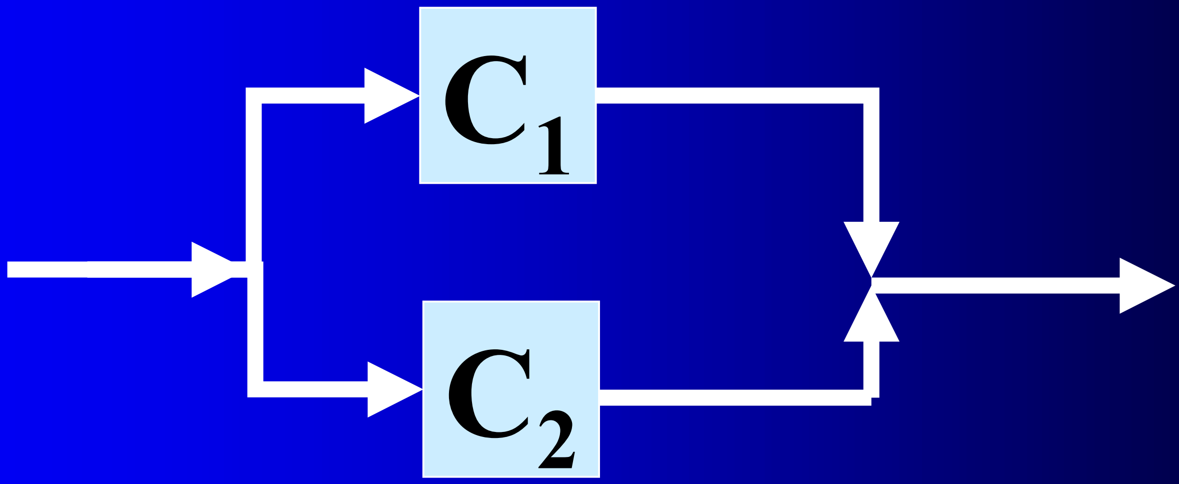
$$S_1(t) = \exp(-1000/1500) = 0.51$$

Y el sistema:

$$S_{sistema}(t) = \exp(-1000/857) = 0.31$$

Sistemas paralelos

- Consta de dos o más componentes en paralelo.
- Funciona mientras un solo componente lo haga.
- Se estropea cuando TODOS los componentes han dejado de funcionar.



Función de Supervivencia

$$P_S(NoFunc) = P_1(NoFunc).P_2(NoFunc)$$

$$P(NoFunc) = 1 - P(Func)$$

$$P(Func.en.t) = S(t)$$

$$`S_S(t) = 1 - (1 - S_1(t))(1 - S_2(t))$$

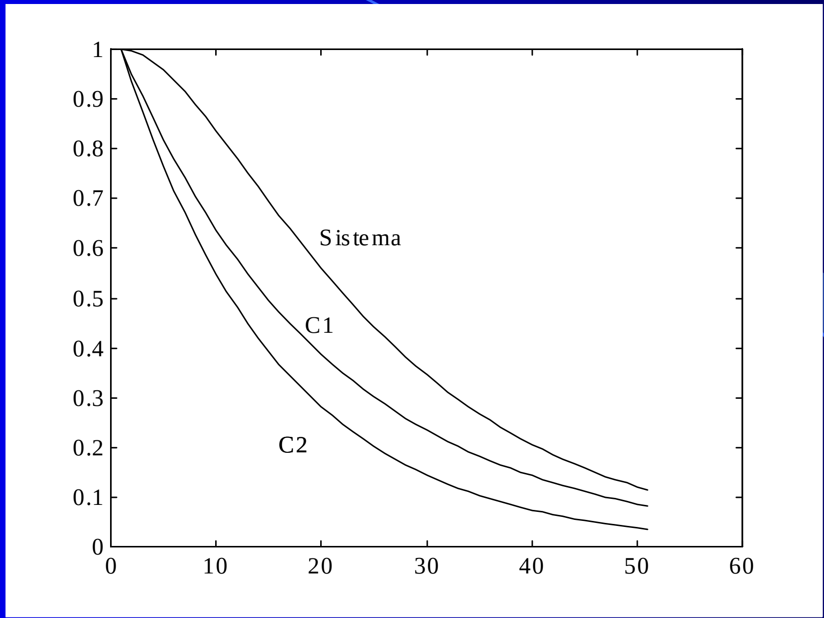
Ejemplo

- Primer componente
- $S_1(t)=\exp(-t/2000)$
- Segundo componente
- $S_2(t)=\exp(-t/1500)$
- Sistema:

$$S_s(t)=1-(1-\exp(-t/2000)).(1-\exp(-t/1500))$$

$$S_s(t)=\exp(-t/2000)+\exp(-t/1500)-\exp(-t/857)$$

$S_s(t)= \exp(-t/2000)+ \exp(-t/1500)- \exp(-t/857)$



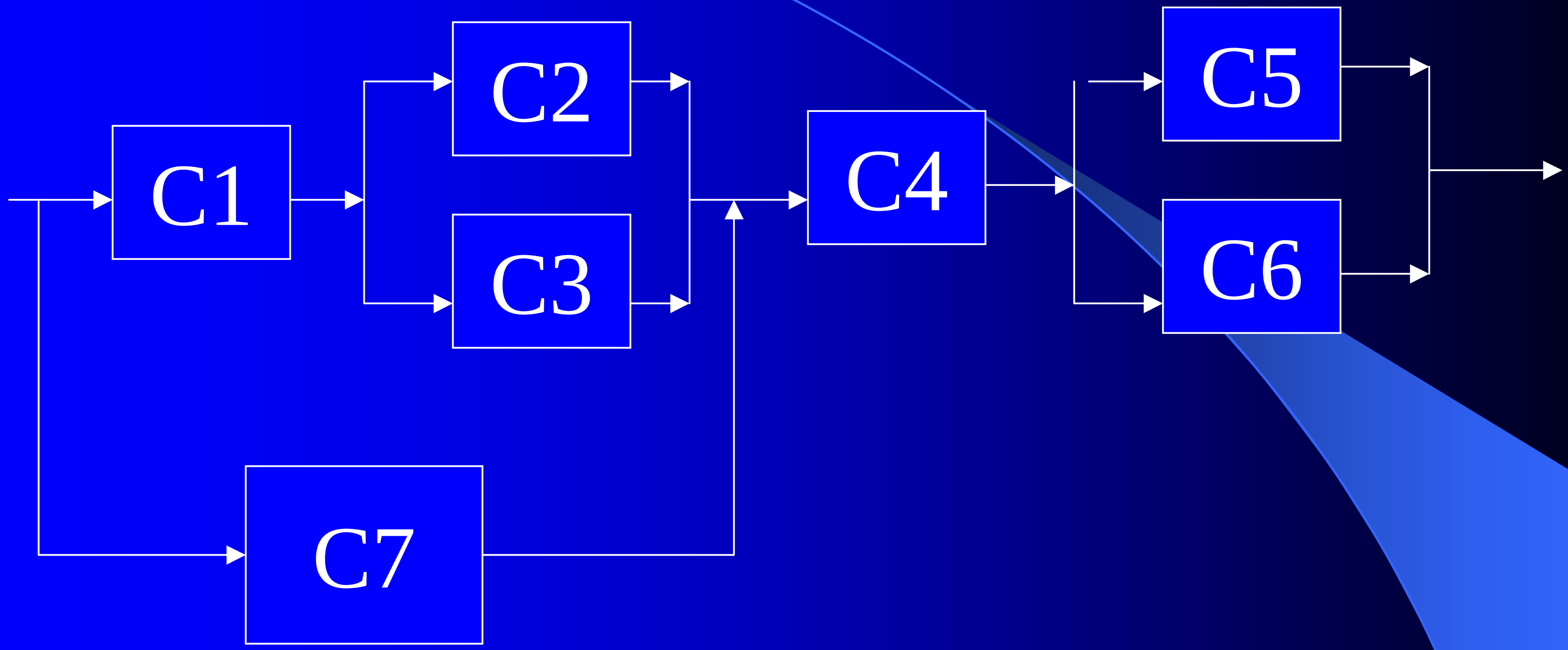
Así, por ejemplo la fiabilidad a las 1000 horas será:

$$S_1(t) = \exp(-1000/2000) = 0.61$$

$$S_1(t) = \exp(-1000/1500) = 0.51$$

$$S_{sistema}(t) = \exp(-1000/2000)+\exp(-1000/1500)-\exp(-1000/857) = 0.8086$$

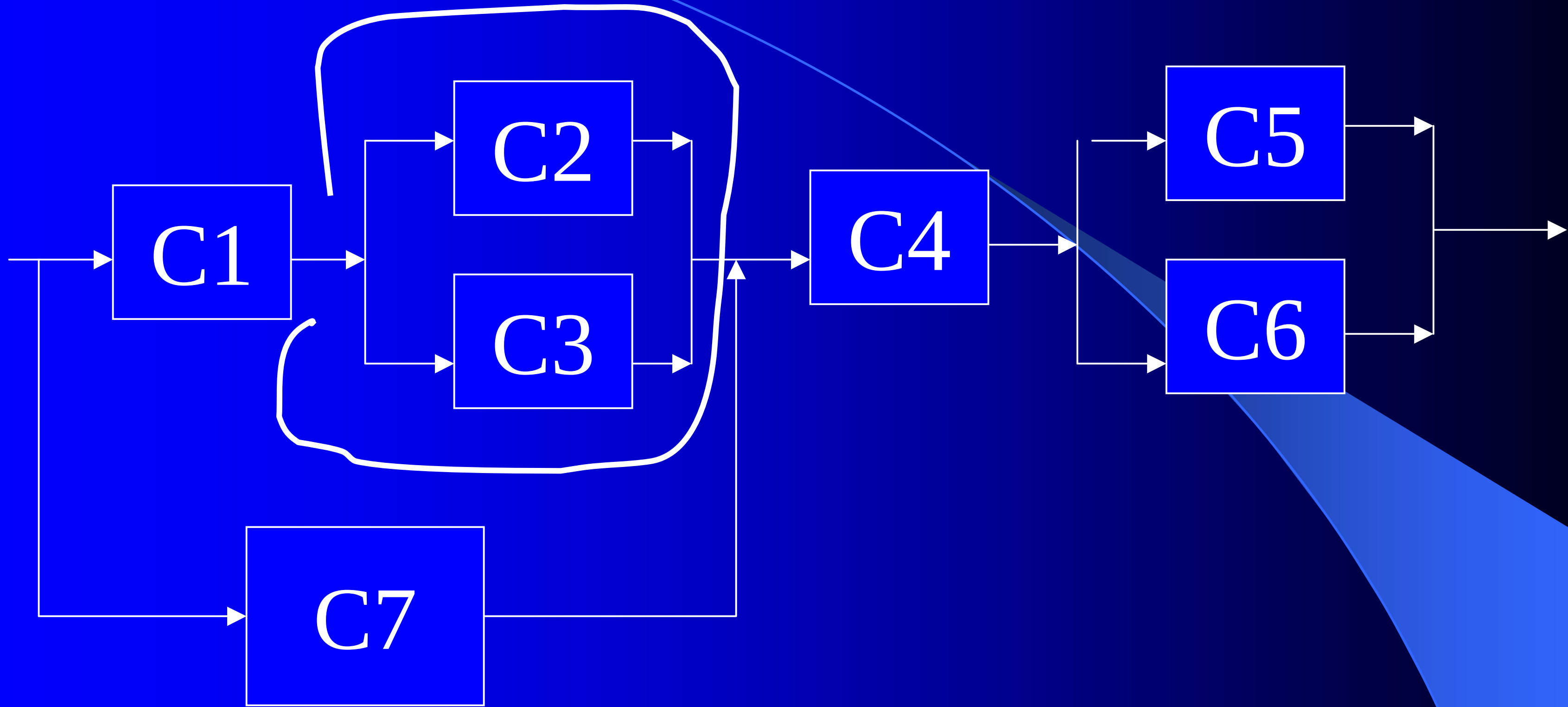
La fiabilidad del sistema paralelo es mayor que la de cualquiera de Sus componentes. En piezas clave se suele redundar un componente



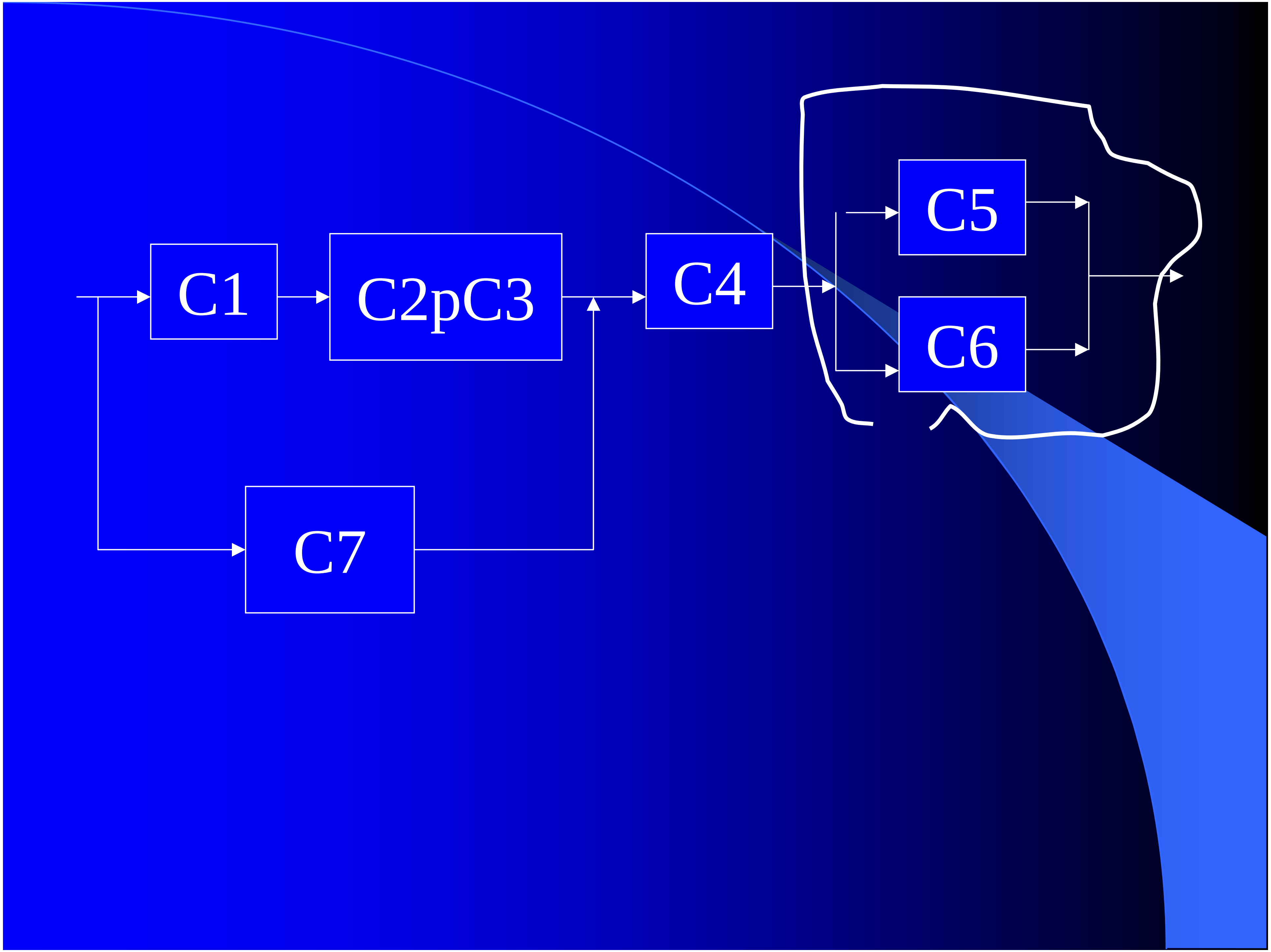
Sistemas complejos

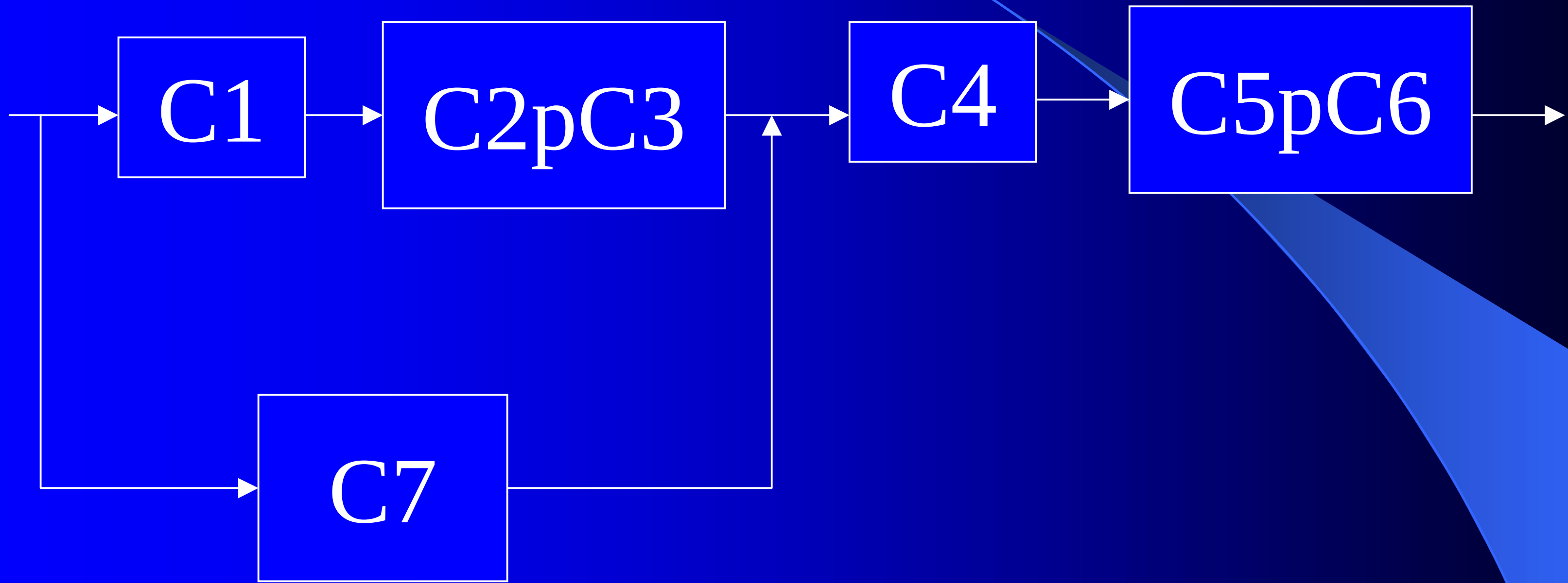
Sistemas complejos

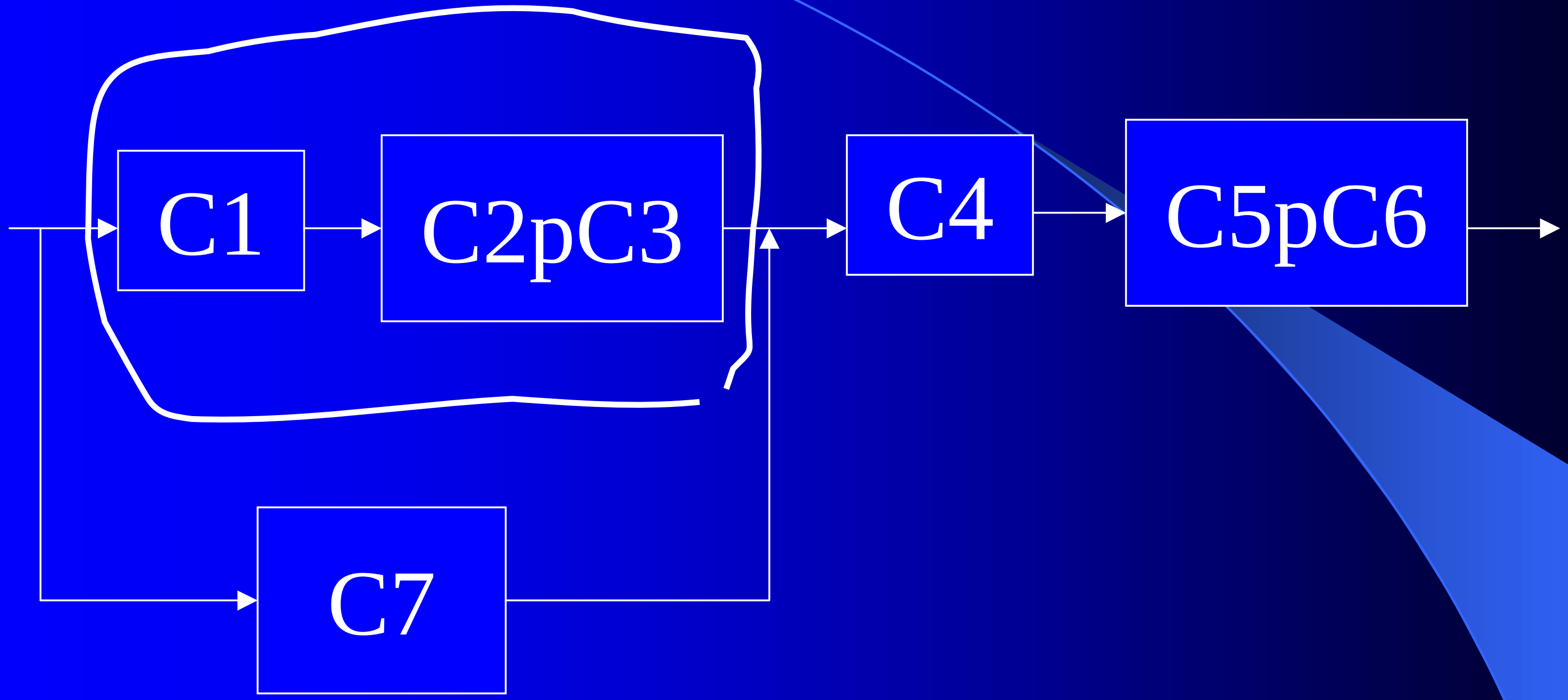
- Hay que ir resolviendo por partes pequeñas
- En etapas
- Esto sólo sirve para pequeños sistemas complejos
- Los grandes sistemas (Una central nuclear) utilizan otros métodos.

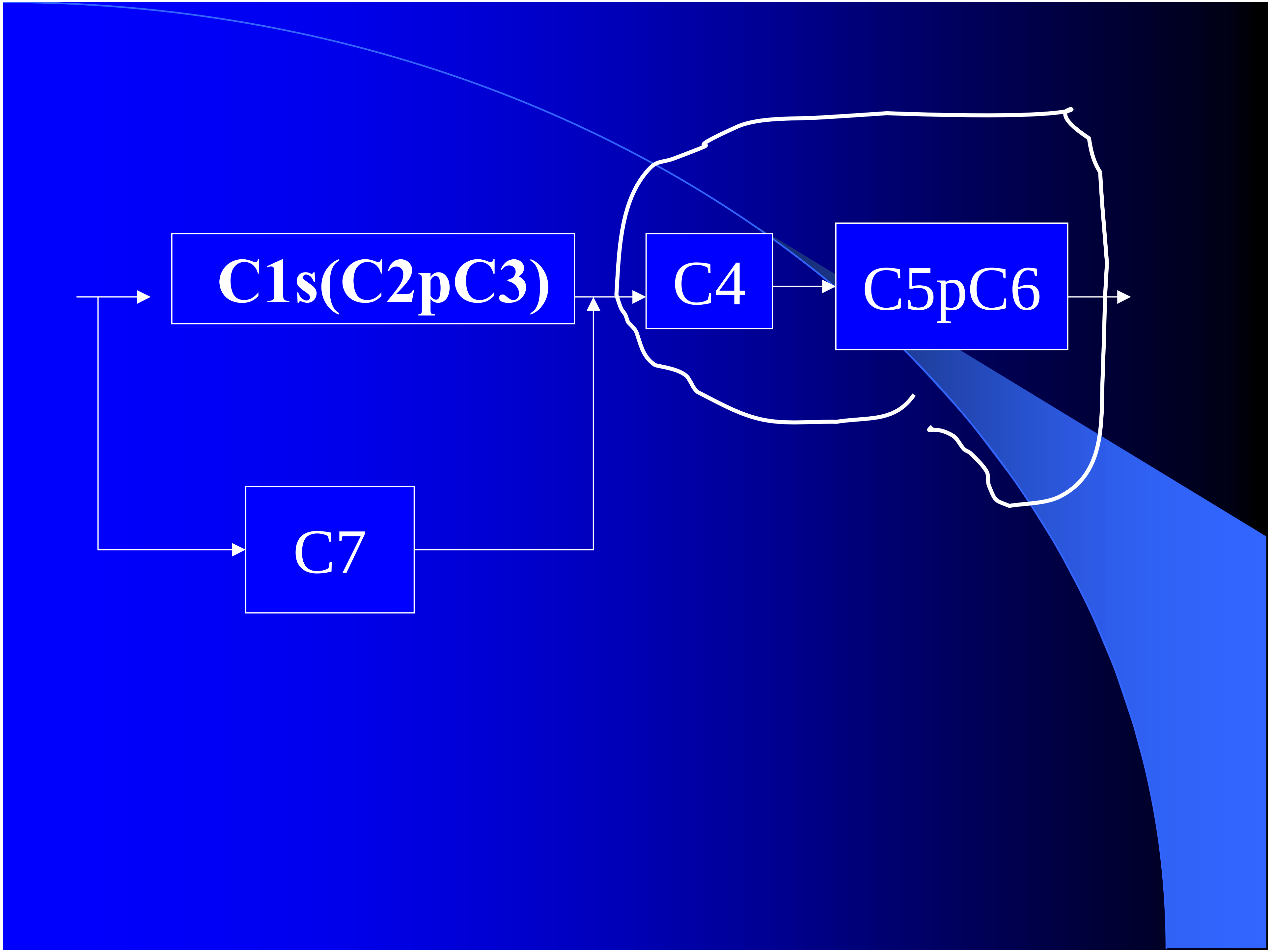


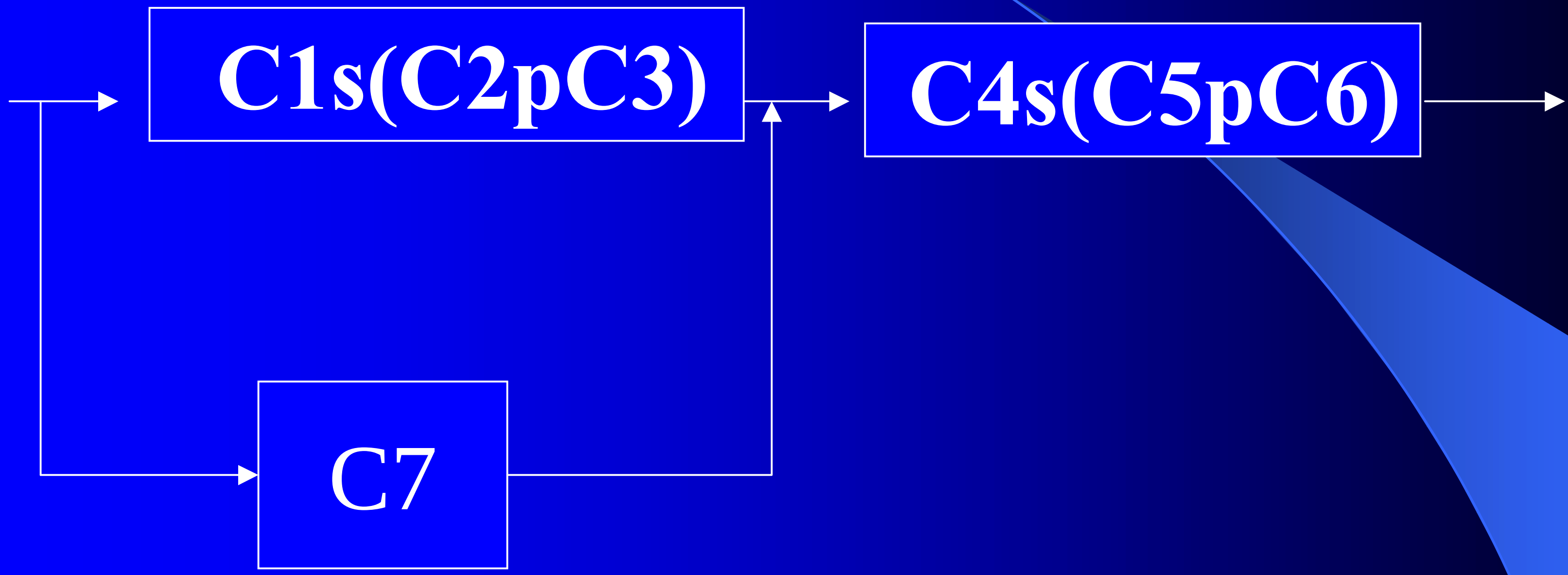
Sistemas complejos

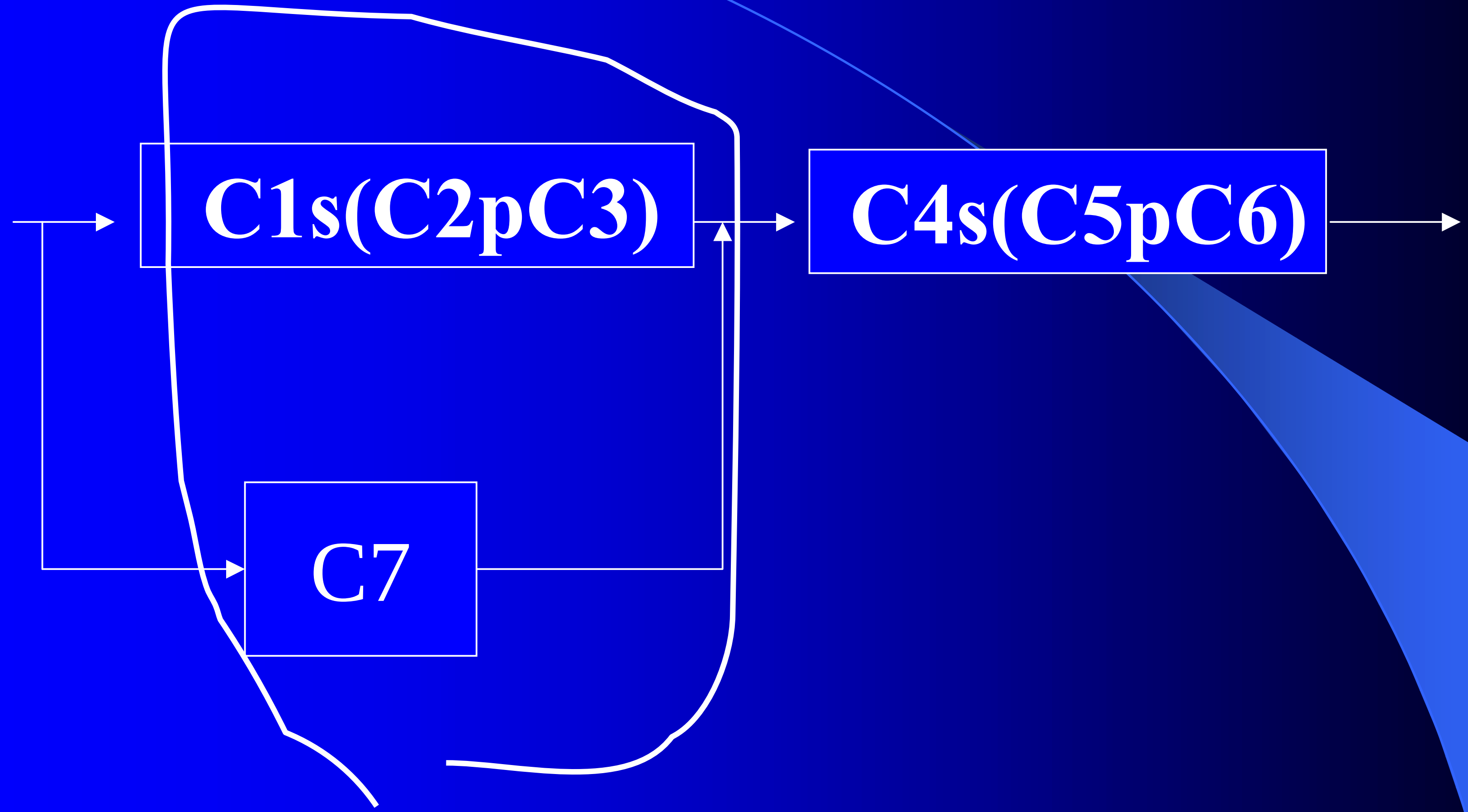

C1
C2pC3
C4
C5
C6
C7


C1
C2pC3
C4
C5pC6
C7


C1
C2pC3
C4
C5pC6
C7


C1s(C2pC3)
C4
C5pC6
C7


C1s(C2pC3)
C4s(C5pC6)
C7


C1s(C2pC3)
C4s(C5pC6)
C7

→ C7s(C1s(C2pC3)) → C4s(C5pC6) →

C7s(C1s(C2pC3))sC4s(C5pC6)

Sistema equivalente

FIN