Distribuciones de Intervalos de tiempo

Jhon Jairo Padilla Aguilar, PhD

Introducción

- Como se mencionó anteriormente, los intervalos de tiempo son variables aleatorias que son de gran importancia en Teletráfico.
- Las distribuciones de estas v.a. son fundamentales para los modelos estudiados.
- La más importante de las distribuciones en Teletráfico es la distribución exponencial, ya que es la base de muchos modelos utilizados por mucho tiempo.
- Se pueden obtener otros tipos de distribuciones mediante la combinación de variables aleatorias exponenciales en serie y en paralelo.

Aplicación de diferentes variables aleatorias en Teletráfico

Distribución	Aplicación
exponencial	Tiempos entre llegadas de llamadas, tiempos de servicio, longitudes de paquetes
Poisson	Número de llegadas que ocurren con distribuciones temporales exponenciales
Paretto	Longitud de páginas web, Longitud de archivos
MMPP (Markov Modulated Poisson Process)	Llegadas de paquetes a un enlace de salida de una red de acceso a una Internet.

Distribución exponencial

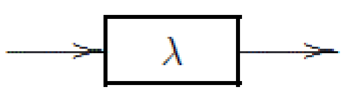
- También se conoce como **distribución exponencial negativa**
- Para modelar los tiempos de la teoría de teletráfico se puede usar cualquier v.a que tenga valores no-negativos para modelar el tiempo de vida.
- La distribución exponencial tiene unas características únicas que la hacen muy apetecida para usos prácticos y analíticos.

Distribución exponencial

- Se caracteriza por un parámetro único: la intensidad o tasa (λ).
- La distribución exponencial tiene la forma:

$$
\begin{aligned}
F(t) &= 1 - e^{-\lambda t}\,, \quad \lambda > 0\,, \quad t \geq 0\,, \\
f(t) &= \lambda\, e^{-\lambda t}\,, \qquad \lambda > 0\,, \quad t \geq 0\,.
\end{aligned}
$$

- Los parámetros que caracterizan la distribución exponencial son:



Representación en los Diagramas de Fase

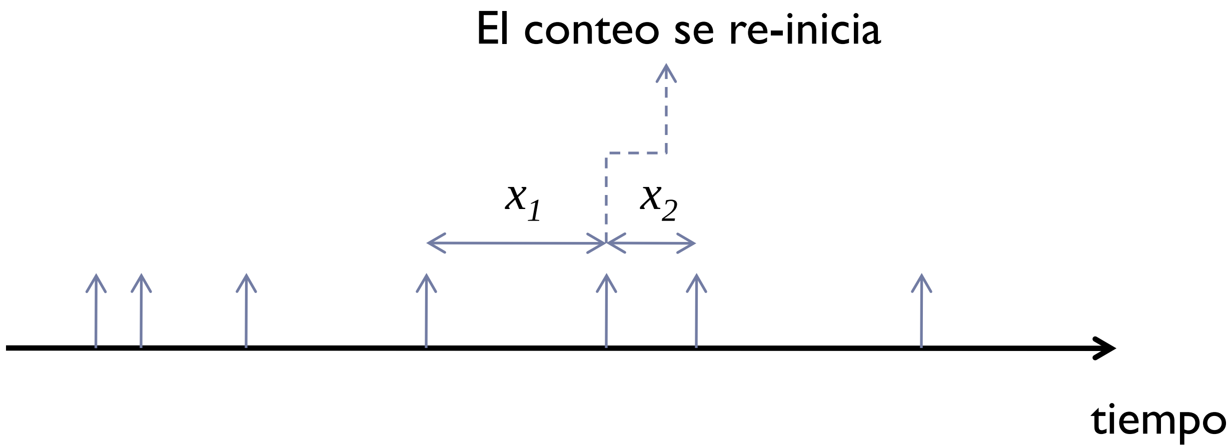
$$
\begin{aligned}
&\text{Mean value } m_1 = \frac{1}{\lambda}\,, \\
&\text{Second moment: } m_2 = \frac{2}{\lambda^2}\,, \\
&\text{Variance: } \sigma^2 = \frac{1}{\lambda^2}\,, \\
&\text{Form factor: } \varepsilon = 2\,,
\end{aligned}
$$

Distribución exponencial

- Propiedades:
  - Es muy apropiada para describir intervalos de tiempo físicos
  - Tiene la propiedad de Falta de Memoria (Ausencia de Memoria)

Propiedad de Falta de Memoria

- Una v.a. X con distribución exponencial también sufre de la propiedad de falta de memoria.
- Esta propiedad significa que la probabilidad de ocurrencia de un evento después de cierto tiempo (o distancia) no tiene en cuenta qué ocurrió antes de iniciar el conteo.



Ejemplo

- El tiempo de vida residual de una conversación telefónica (lo que queda de ella, dado que ya transcurrió un tiempo x), puede suponerse independiente de la duración actual de la llamada.
- Por tanto, puede suponerse que tiene una distribución con falta de memoria tal como la distribución exponencial:

$$
\begin{aligned}
f(t+x|x) &= \frac{\lambda \mathrm{e}^{-(t+x)\lambda}}{\mathrm{e}^{-\lambda x}} \\
&= \lambda \mathrm{e}^{-\lambda t} \\
&= f(t)\,.
\end{aligned}
$$

Mínimo de dos v.a. exponenciales

- Si asumimos que dos variables aleatorias X_1 y X_2 son mutuamente independientes y exponencialmente distribuidas con intensidades λ_1 y λ_2 respectivamente, se puede definir una nueva v.a. X como:

$$X = min\{X_1, X_2\}$$

- Esta función de distribución tendrá la forma:

$$p\{X \leq t\} = 1 - e^{-(\lambda_1 + \lambda_2)t} .$$

- Es decir, es una distribución exponencial con intensidad $(\lambda_1 + \lambda_2)$

Mínimo de dos v.a. exponenciales (II)

- Suponiendo que $X_1 < X_2$, la probabilidad de ocurrencia de X_1 primero que X_2 (y que X_2 ocurra después de un tiempo t), dado que ya pasó un tiempo t es:

$$
\begin{aligned}
p\{X_1 < X_2 | \, t\} &= \frac{P\{t < X_1 \leq t + \mathrm{d}t\} \cdot P\{X_2 > t\}}{P\{t < X \leq t + \mathrm{d}t\}} \\
&= \frac{\lambda_1 \, \mathrm{e}^{-\lambda_1 t} \, \mathrm{d}t \cdot \mathrm{e}^{-\lambda_2 t}}{(\lambda_1 + \lambda_2) \, \mathrm{e}^{-(\lambda_1 + \lambda_2) t} \, \mathrm{d}t} \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2},
\end{aligned}
$$

- En este caso, esta probabilidad es independiente de t. Por tanto, no necesitamos integrar con respecto a t.

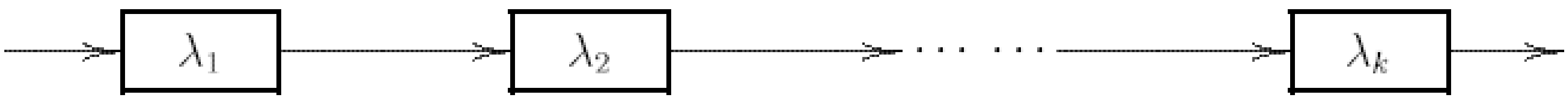
Variables aleatorias derivadas de la v.a. exponencial

Combinación de v.a. exponenciales

- Si una v.a. exponencial no puede describir los intervalos de tiempo con suficiente detalle, se puede usar una combinación de dos o más v.a. exponenciales.
- Conny Palm introdujo dos clases de distribuciones: steep (con pendiente) y flat (planas)
- Distribuciones con pendiente:
  - Corresponden a un conjunto de intervalos de tiempo exponencialmente distribuidos e independientes en serie.
- Distribuciones Planas:
  - Corresponden a un conjunto de intervalos de tiempo exponencialmente distribuidos e independientes en paralelo.
- Además, mediante la combinación de distribuciones planas y con pendiente, se puede obtener arbitrariamente una buena aproximación para cualquier distribución.

Distribuciones con Pendiente (Steep)

Diagrama de Fase:



- Se obtienen combinando k distribuciones exponenciales en serie.
- Se conocen como distribuciones hipo-exponenciales o Distribuciones Erlang generalizadas.
- Tienen un factor de forma $1<\varepsilon<2$
- Si todas las k distribuciones son idénticas ($\lambda_i=\lambda$), se obtiene una *Distribución Erlang-k*.
- Se les llama distribuciones con pendiente porque van de cero a uno más rápido que una distribución exponencial.

Distribución Erlang-k

- La distribución de Erlang-k tiene las siguientes expresiones para la f.d.p. y la Distribución acumulada respectivamente:

$$f(t) = \frac{(\lambda t)^{k-1}}{(k-1)!} \cdot \lambda \cdot e^{-\lambda t}, \quad \lambda > 0, \quad t \geq 0, \quad k = 1, 2, \ldots .$$

$$F(t) = \sum_{j=k}^{\infty} \frac{(\lambda t)^j}{j!} \cdot e^{-\lambda t}$$

$$= 1 - \sum_{j=0}^{k-1} \frac{(\lambda t)^j}{j!} \cdot e^{-\lambda t}$$

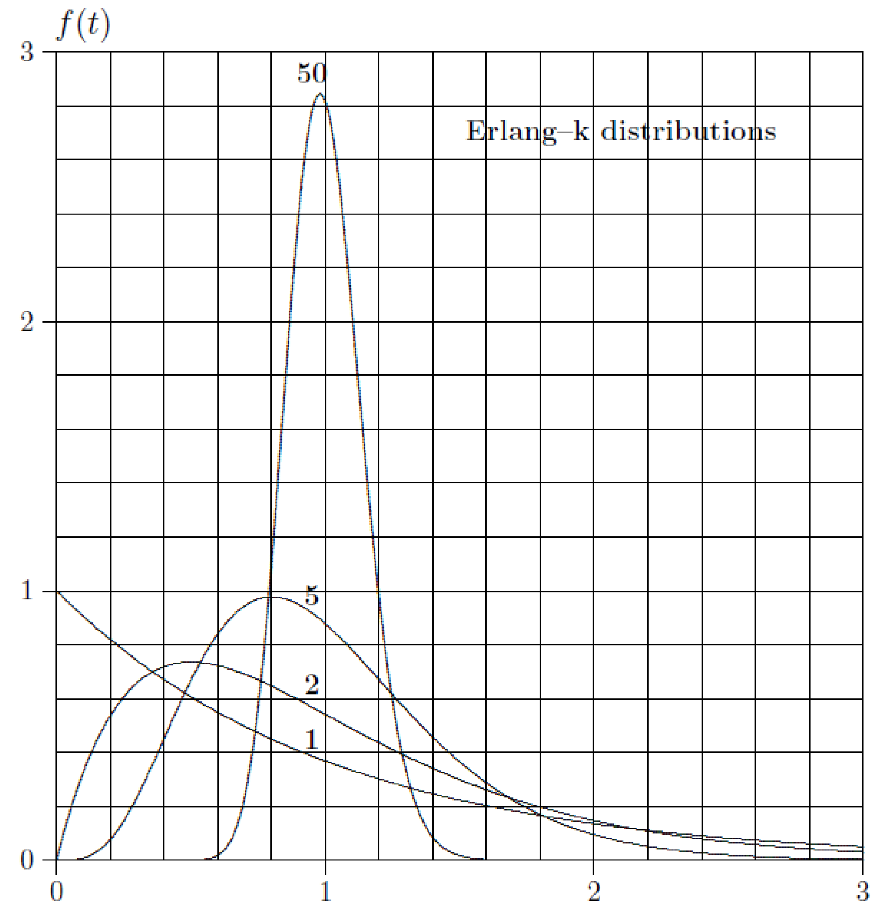
Distribución Erlang-K

- La caracterizan los siguientes momentos:

$$
\begin{aligned}
m &= \frac{k}{\lambda}\,, \\
\sigma^2 &= \frac{k}{\lambda^2}\,, \\
\varepsilon &= 1 + \frac{\sigma^2}{m^2} = 1 + \frac{1}{k}\,,
\end{aligned}
$$

- Y el i-ésimo momento no central es:

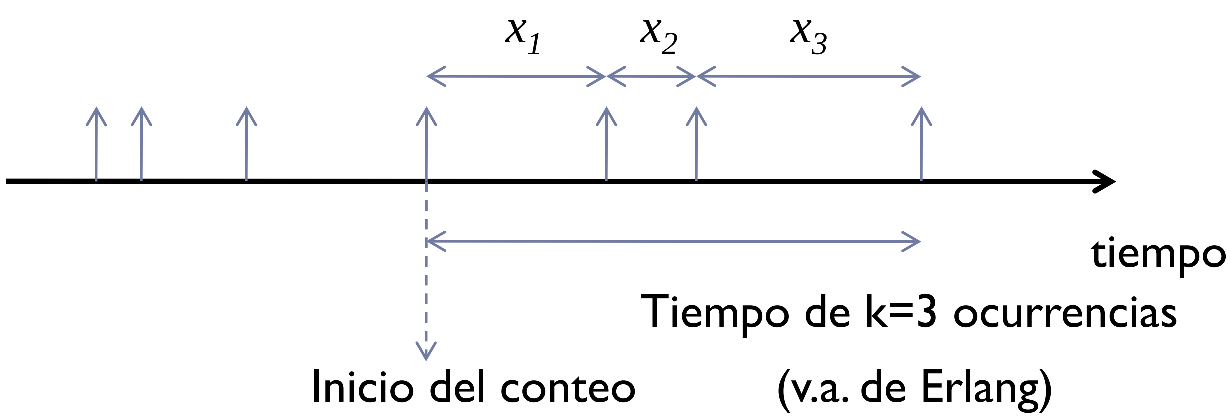
$$m_i = \frac{(i+k-1)!}{(k-1)!} \cdot \left(\frac{1}{\lambda}\right)^i .$$



Distribución de Erlang-k para diferentes valores de k (con λ=1). Para k=1 se comporta como una distribución exponencial

Aplicaciones de Erlang-K

- La distribución de Erlang-k describe el tiempo (o longitud) hasta que suceden k ocurrencias en un proceso de Poisson con media λ.



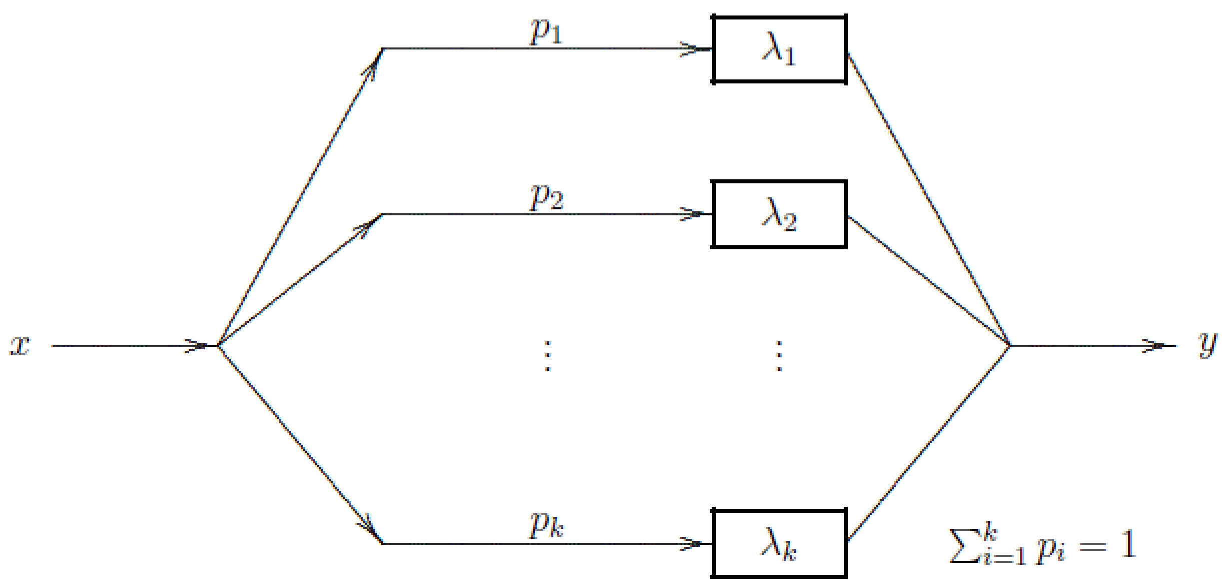
- Ejemplos:
  - Si tenemos tiempos con distribución exponencial entre llegadas de llamadas a una central telefónica. La distribución Erlang-k mediría el tiempo necesario para que lleguen k llamadas.
  - Si tenemos longitudes de paquetes que llegan a una cola con una distribución exponencial, la distribución Erlang-k mediría el tiempo necesario para atender k paquetes (suponiendo un tiempo de atención constante en bits/seg).

Distribuciones Planas

- La distribución plana general es una suma ponderada de distribuciones exponenciales y tiene la Distribución acumulada:

$$F(t) = \int_0^\infty \left(1 - e^{-\lambda t}\right) dW(\lambda), \qquad \lambda > 0, \quad t \geq 0,$$

- La función de peso, $W(\lambda)$, podría ser continua o discreta.
- La distribución plana se obtiene al combinar k distribuciones en paralelo y elegir una rama i con una probabilidad pi.
- Cuando $W(\lambda)$ es discreta, el resultado es una distribución hiper-exponencial.



Distribución Hiper-exponencial

- Supóngase que las tasas de las distribuciones exponenciales son respectivamente: $\lambda_1, \lambda_2, \ \ldots \ , \lambda_k \, ,$

- Y que *W(λ)* tiene los incrementos positivos: $p_1, p_2, \ \ldots \ , p_k \, ,$

- Donde, $\sum_{i=1}^{k} p_i = 1 \, .$

- Para otros valores *W(λ)* es constante. La distribución acumulada será:

$$F(t) = 1 - \sum_{i=1}^{k} p_i \cdot \mathrm{e}^{-\lambda_i t} \, , \quad t \geq 0 \, .$$

Distribución Hiper-exponencial

- La media y el factor de forma serán:

$$m_1 = \sum_{i=1}^{k} \frac{p_i}{\lambda_i},$$

$$\varepsilon = \left\{\sum_{i=1}^{k} p_i \frac{2}{\lambda_i^2}\right\} \Big/ \left\{\sum_{i=1}^{k} \frac{p_i}{\lambda_i}\right\}^2 \geq 2.$$

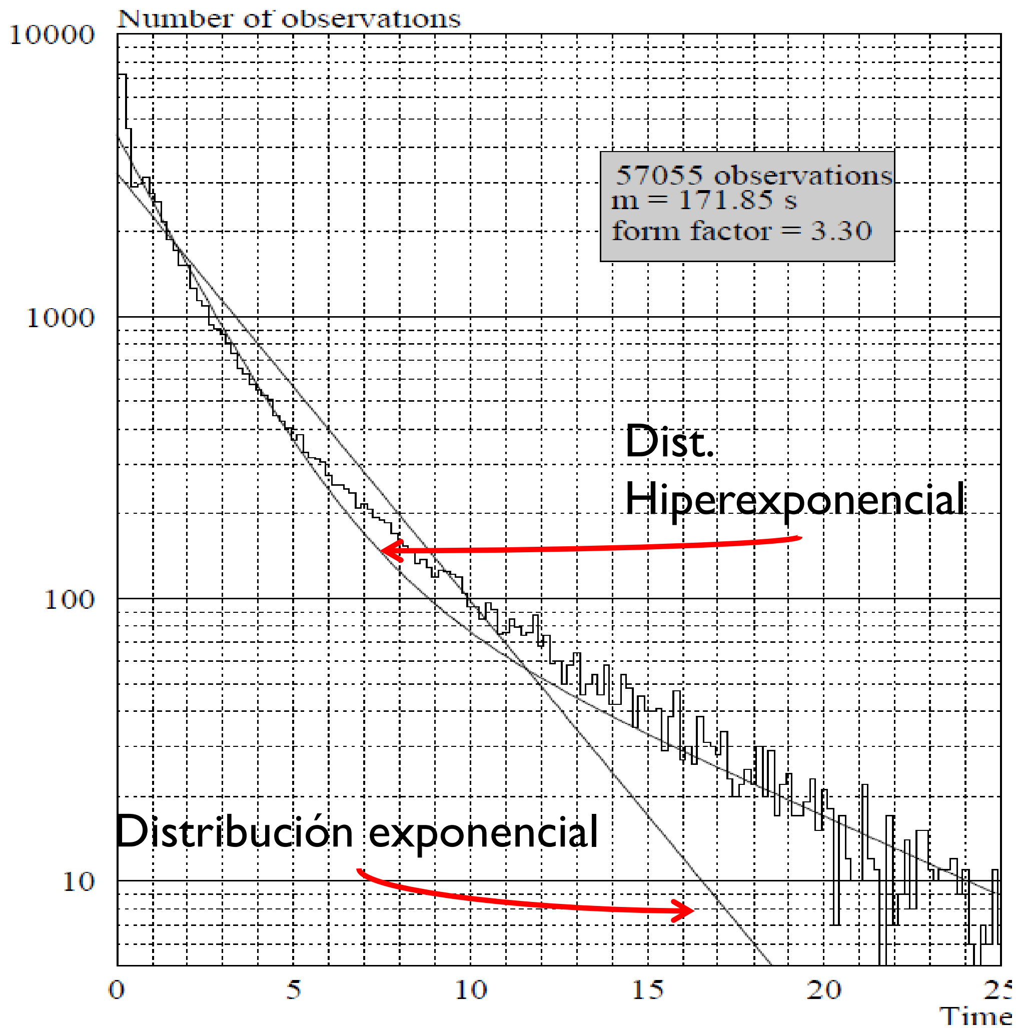
- Si $k=1$ ó todas las λ_i son iguales, se tendrá una distribución exponencial.
- Las distribuciones planas se llaman así porque la variación de cero a uno es más lenta que la distribución exponencial.

Distribución hiper-exponencial

- En la práctica es difícil estimar más de uno o dos parámetros.
- El caso más importante es para $n=2$ ($p_1=p$, $p_2=1-p$). Por tanto,

$$F(t) = 1 - p \cdot \mathrm{e}^{-\lambda_1 t} - (1-p) \cdot \mathrm{e}^{-\lambda_2 t}\,.$$

Aplicación de la distribución hiper-exponencial

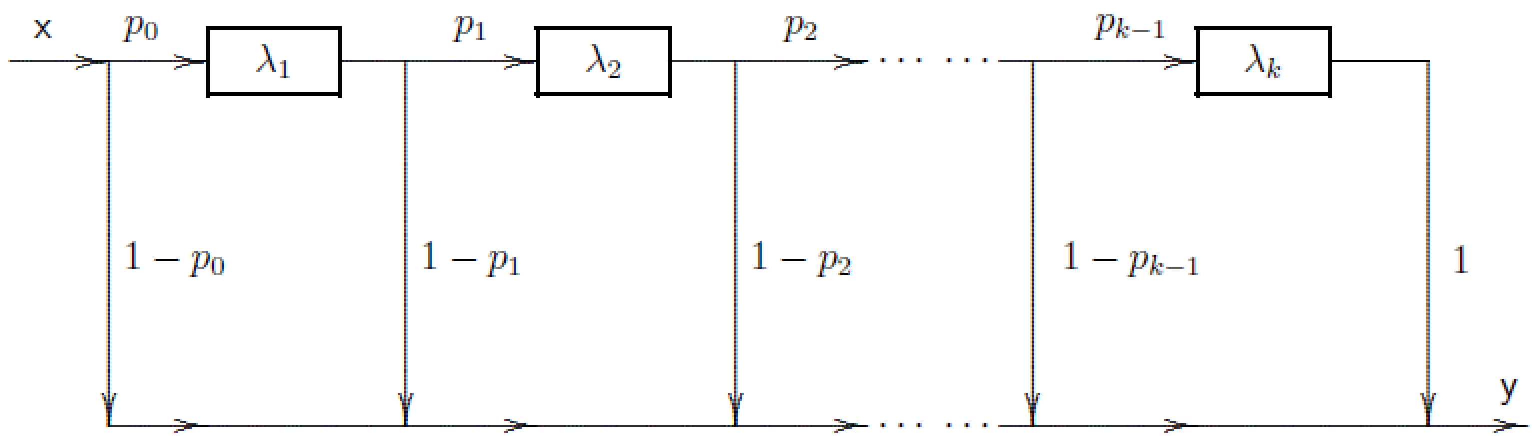


- La gráfica muestra mediciones de los tiempos de servicio en líneas de una central telefónica local durante la hora pico.

Distribuciones de Cox

- Se obtienen combinando distribuciones con pendiente y planas.
- Es una clase de distribuciones general que puede ser descrita mediante combinaciones de distribuciones exponenciales en serie y en paralelo.
- Aquí se estudiará un caso especial conocido como Erlang con ramificaciones (Branching Erlang)

Distribución Erlang con ramificaciones



- La media es,

$$m_1 = \sum_{i=1}^{k} q_i \left(1 - p_i\right) \left\{ \sum_{j=1}^{i} \frac{1}{\lambda_j} \right\} ,$$

- Donde $q_i = p_0 \cdot p_1 \cdot p_2 \cdot \cdots \cdot p_{i-1}$.
- El término $q_i(1-p_i)$ corresponde a la probabilidad de ramificarse, y es la probabilidad de saltar afuera después de dejar la fase i.

Distribución Erlang con ramificaciones

- La media es,

$$m_1 = \sum_{i=1}^{k} \frac{q_i}{\lambda_i} = \sum_{i=1}^{k} m_{1,i}\,,$$

- El segundo momento es,

$$m_2 = 2 \cdot \sum_{i=1}^{k} \left\{ \left(\sum_{j=1}^{i} \frac{1}{\lambda_j} \right) \cdot \frac{q_i}{\lambda_i} \right\}$$

- La varianza será,

$$\sigma^2 = m_2 - m_1^2\,.$$

Propiedades de la Distribución de Cox

- La suma de dos v.a. de Cox dan como resultado otra distribución de Cox.
- La f.d.p. de una Distribución de Cox se puede escribir como una sumatoria de v.a. exponenciales:

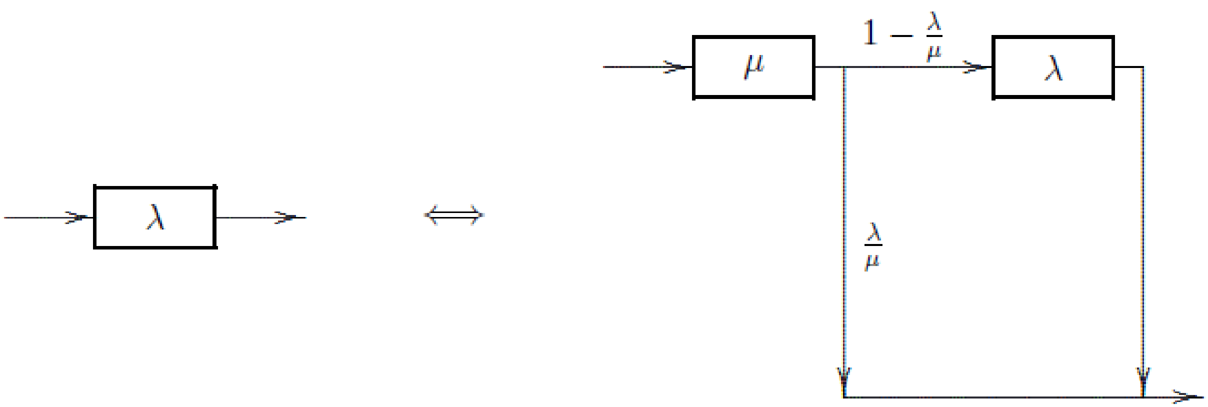
$$1 - F(t) = \sum_{i=1}^{k} c_i \cdot \mathrm{e}^{-\lambda_i t} \, ,$$

- Donde

$$0 \leq \sum_{i=1}^{k} c_i \leq 1 \, , \qquad -\infty < c_i < +\infty \, .$$

Principios de descomposición

- Los diagramas de fase son una herramienta útil para analizar las distribuciones de Cox a partir de las propiedades de la distribución exponencial.
- Teorema 1:
  - Una distribución exponencial con intensidad λ puede ser descompuesta en una distribución de Cox de dos fases, donde la primera fase tiene una intensidad $\mu>\lambda$ y la segunda fase tiene la intensidad original λ

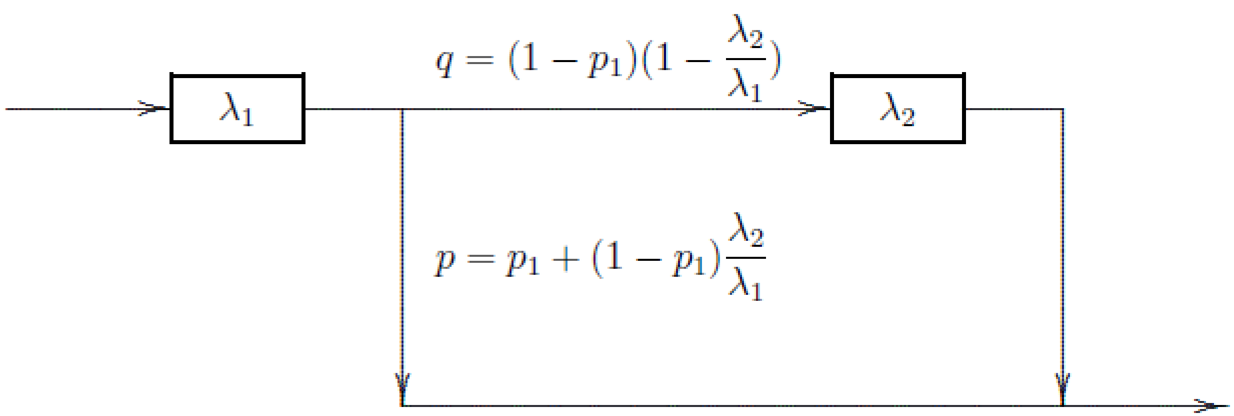


Teorema 2:

Las fases en cualquier distribución de Cox pueden ser ordenadas como $\lambda_i \geq \lambda_{i+1}$

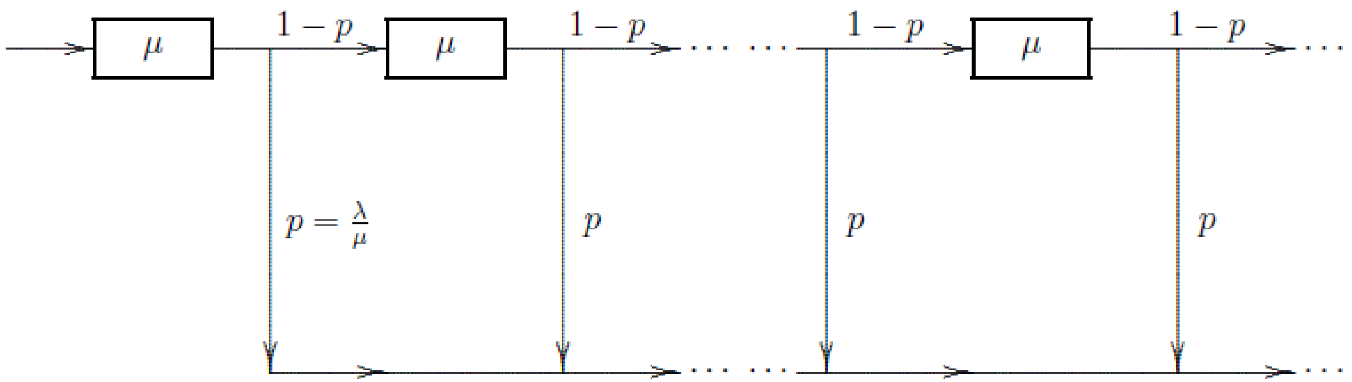
Cox como base de otras distribuciones

- En consecuencia, una **distribución Hiper-exponencial** con *l* fases es equivalente a una distribución de Cox con el mismo número de fase $(\lambda_1 > \lambda_2,\ p_2 = 1 - p_1)$
- Caso para *l=2*:



Cox como base de otras Distribuciones

- Una **distribución exponencial** es equivalente a una distribución de Cox homogénea (la misma intensidad en todas las fases) con intensidad μ y un número de fases infinito. Las probabilidades de ramificación son constantes.



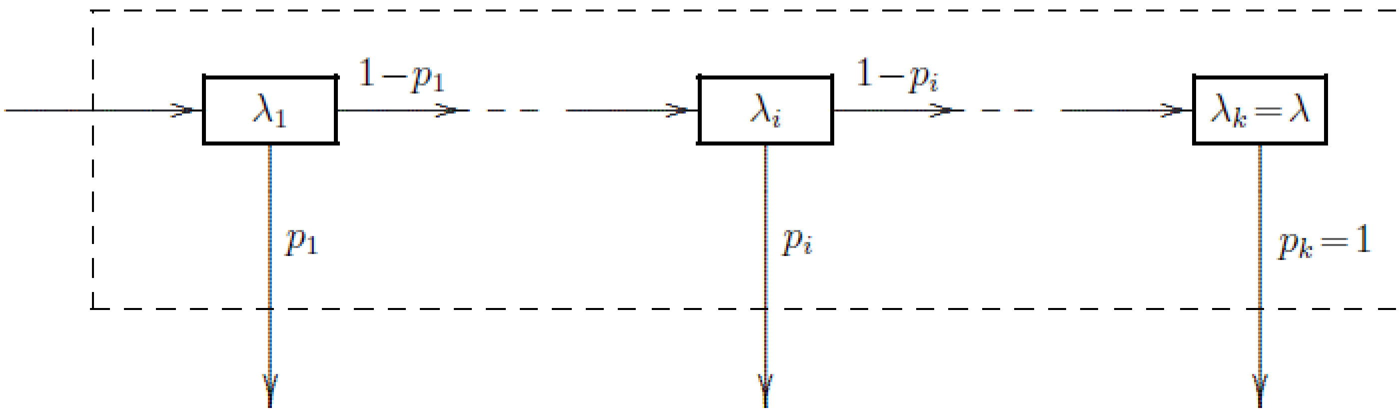
- Por tanto, una distribución exponencial se puede ver como una distribución compuesta de distribuciones Erlang-k homogéneas con tasas $\mu>\lambda$, donde los factores de peso siguen una distribución geométrica (cociente $p=\lambda/\mu$).

Cox como base de otras distribuciones

- Una **distribución exponencial** también puede verse como la distribución de fases de la figura. Esto se cumple cuando,

$$p_i \cdot \lambda_i = \lambda. \quad \lambda_i \geq \lambda \qquad 0 < p_i \leq 1$$

- Y el número de fases es finito e igual a k.



Importancia de la distribución de Cox

- Ha atraído la atención durante años recientes.
- Es de gran importancia porque posee las siguientes propiedades:
  - La distribución de Cox puede ser analizada usando el método de fases
  - Cualquier distribución puede ser aproximada de una forma bastante buena usando una distribución de Cox.
  - Si una propiedad es válida para una distribución de Cox, entonces es válida para cualquier distribución de interés práctico.
- En la práctica:
  - En general, si suponemos que hay 2k parámetros en un problema estadístico no resuelto. Normalmente, podemos elegir una distribución de Cox especial (p.ej. Erlang-k o Hiper-exponencial) y aproximar el primer momento.

Notación para las principales distribuciones

M	$\sim$	Exponential distribution (<u>M</u>arkov),
E_k	$\sim$	Erlang-k distribution,
H_n	$\sim$	Hyper-exponential distribution of order n,
D	$\sim$	Constant (<u>D</u>eterministic),
Cox	$\sim$	Cox distribution,
G	$\sim$	General = arbitrary distribution.

Otras distribuciones de tiempo

- Distribución gamma:
  - Se obtiene al suponer que el parámetro k de una distribución Erlang-k toma valores reales no negativos:

$$f(t) = \frac{1}{\Gamma(k)}(\lambda t)^{k-1} \cdot \mathrm{e}^{-\lambda t} \cdot \lambda\,, \qquad \lambda > 0\,, \quad t \geq 0\,.$$

$$m = \frac{k}{\lambda}\,,$$

$$\sigma^2 = \frac{k}{\lambda^2}\,,$$

Otras distribuciones de tiempo

- Distribución Weibull:
  - Esta distribución tiene una intensidad de muerte (tasa de servicio), μ, dependiente del tiempo.
  - Tiene su origen en la teoría de la confiabilidad.
  - Para k=1, se obtiene la distribución exponencial.

$$F(t) = 1 - \mathrm{e}^{-(\lambda t)^k}, \qquad t \geq 0, \quad k > 0, \quad \lambda > 0.$$

$$\begin{aligned} \frac{\mathrm{d}F(t)}{1 - F(t)} = \mu(t) &= \frac{\lambda \mathrm{e}^{-(\lambda t)^k} \cdot k\,(\lambda t)^{k-1}\,\mathrm{d}t}{\mathrm{e}^{-(\lambda t)^k}} \\ &= \lambda k\,(\lambda t)^{k-1}. \end{aligned}$$

Otras distribuciones de tiempo

- Distribución de Pareto:

$$F(t) = 1 - (1 + \eta_0 t)^{-\left(1+\frac{\lambda}{\eta_0}\right)} .$$

$$m_1 = \frac{1}{\lambda},$$

$$\varepsilon = \frac{2\lambda}{\lambda - \eta_0}, \qquad 0 < \eta_0 < \lambda .$$

- Note que la varianza no existe para $\lambda \leq \eta_0$.
- Si se hace $\eta_0 \to 0$, se convierte en una distribución exponencial.
- Si la intensidad de un proceso de Poisson tiene una distribución Gamma, entonces los tiempos entre llegadas tienen una distribución de Pareto.

Distribuciones de cola pesada

- En sistemas telefónicos rara vez se tienen mediciones con factores de forma mayores que 6.
- En tráfico de datos, se obtienen mediciones con factores de forma mayores que 100.
- Esto indica que la variación de las mediciones es muy alta. Este es el caso de las transmisiones de datos, donde podemos enviar unos pocos caracteres o una gran cantidad de datos.
- Para describir estos datos, se utilizan distribuciones de cola pesada (heavy-tailed).
- Una distribución se considera de cola pesada si la cola de la función de distribución se comporta como una ley de potencias, es decir, $1 - F(t) \approx t^{-\alpha}, \; 0 < \alpha \leq 2.$

Distribuciones de cola pesada

- La distribución de Pareto es de cola pesada en sentido estricto.
- Ademas, se consideran distribuciones de cola pesada, aquellas cuya cola es más pesada que la de la distribución exponencial.
- Ejemplos:
  - Hiper-exponencial
  - Weibull
  - Lognormal
- Recientemente, medidas más extensas han permitido modelar el tráfico de datos mediante modelos auto-similares.

Distribuciones Discretas

Proceso de Bernoulli

- El experimento consiste de n ensayos que se repiten
- Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso (éxito y fracaso son sólo etiquetas, no tienen el significado de que algo esté bien o mal)
- La probabilidad de un éxito se denota como p y permanece constante de un ensayo a otro
- Los ensayos que se repiten son independientes

Variable aleatoria binomial

- El número de éxitos de n experimentos de Bernoulli se denomina **variable aleatoria binomial**.
- La distribución de probabilidad de esta variable aleatoria discreta se llama **distribución binomial**.
- Distribución binomial: $b(x; n,p)$. Depende del número de ensayos y de la probabilidad de éxito en un ensayo dado.

Distribución binomial

- Un experimento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q=1-p$. Entonces, la distribución de probabilidad de la v.a. binomial X, que es el número de éxitos en n ensayos independientes es,

$$b(x; n, p) = b(x) = \binom{n}{x} p^x q^{n-x} \quad , x=0,1,2,\ldots, n$$

Ejemplo

- Las posibilidades de que un bit transmitido a través de un canal de transmisión digital se reciba con error es de 0.1. Suponga además que los ensayos de transmisión son independientes. Sea X el número de bits con error en los siguientes 4 bits transmitidos. Determine la probabilidad de que lleguen dos bits con error.
- Solución:

p=0.1

n=4

x=2

$$b(x;n,p) = b(x) = \binom{n}{x} p^x q^{n-x}$$

$$b(2) = \binom{4}{2} (0.1)^2 (0.9)^{4-2} = 0,0486$$

Media y Varianza

- La media y la varianza de la distribución binomial b(x;n,p) son

$$\mu = np$$

$$\sigma^2 = npq$$

- Ejemplo:
  - Para el ejemplo de la transmisión de 4 bits con p=0.1, se tiene

$$\mu = (4)\cdot(0,1) = 0,4$$

$$\sigma^2 = (4)\cdot(0.1)\cdot(0.9) = 0,36$$

Distribución Geométrica

- En una serie de ensayos de Bernoulli independientes, con probabilidad constante p de un éxito, sea que la variable aleatoria X denote el número de ensayos hasta el primer éxito. Entonces X tiene una distribución geométrica con parámetro p y

$$f(x) = (1 - p)^{x-1} p \qquad , x = 1, 2, \ldots$$

Ejemplo

- La probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es de 0.1. Suponga que las transmisiones son eventos independientes y sea que la v.a. X denote el número de bits transmitidos hasta el primer error. Calcule la probabilidad de que se transmitan 4 bits correctamente y el quinto bit esté errado.
- *Solución:*

  X=5

  p=0.1

  q=1-0,1=0,9

$$P(X=5)=(0,9)^4(0,1)=0,066$$

Media y varianza de una v.a. con distribución geométrica

- Si X es una v.a. geométrica con parámetro p, entonces la media y la varianza de X son

$$\mu = E(X) = \frac{1}{p}$$

$$\sigma^2 = \frac{(1-p)}{p^2}$$

Ejemplo

- En el ejemplo de la transmisión de bits con probabilidad 0.1 de que llegue un bit errado, calcule el número promedio de bits transmitidos incluyendo el primer bit errado. Calcule la desviación estándar de esta misma variable.
- *Solución:*

p=0,1

$$\mu = E(X) = \frac{1}{0,1} = 10$$

$$\sigma = \left[\frac{(1-p)}{p^2}\right]^{1/2} = \left[\frac{(1-0,1)}{0,01^2}\right]^{1/2} = 9,49$$

Distribución Binomial Negativa

- En una serie de ensayos de Bernoulli independientes, con probabilidad constante p de éxito, sea que la v.a. X denote el número de ensayos hasta que ocurran r éxitos. Entonces X tiene una distribución binomial negativa donde

- Para $x=r, r+1, r+2, \ldots$

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^{r}$$

Aclaraciones

- Debido a que se necesitan al menos r ensayos para obtener r éxitos, el rango de X es de r a infinito.
- Si r=1, una v.a. binomial negativa se convierte en una v.a. geométrica

Ejemplo

- Suponga que la probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es 0,1. Suponga que las transmisiones son eventos independientes, y sea que la v.a. X denote el número de bits transmitidos hasta que se transmite el cuarto bit errado. Calcule la probabilidad de que el décimo bit sea el cuarto bit errado.
- Solución:

r=4 éxitos

x=10 ensayos

$$f(x)=\binom{10-1}{4-1}(1-0,1)^{10-4}0,1^{4}$$

$$f(x)=\binom{9}{3}(0,9)^{6}0,1^{4}=\boxed{\binom{9}{3}(0,9)^{6}0,1^{3}}0,1^{1}$$

Probabilidad de que ocurran 3 errores en los nueve primeros ensayos (Distribución binomial)

Prob. Del cuarto bit errado

Media y Varianza

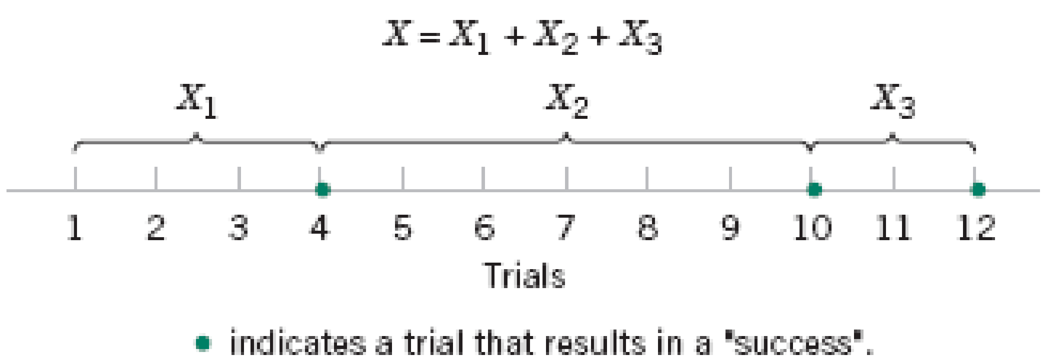
- Si X es una v.a. aleatoria binomial negativa con parámetros p y r, entonces la media y varianza de X son,

$$\mu = E(X) = \frac{r}{p}$$

$$\sigma^2 = \frac{r(1-p)}{p^2}$$

Relación entre las distribuciones geométrica y binomial negativa

- Se dice que una v.a. geométrica no tiene memoria.
- La **propiedad de falta de memoria** significa que cada vez que se obtiene un éxito se re-inicia el conteo de los resultados de los ensayos hasta el siguiente éxito.
- Sea X_1 el número de ensayos hasta el primer éxito, X_2 el número de ensayos hasta el segundo éxito, etc. Entonces, para obtener r éxitos $X=X_1+X_2+\ldots+X_r$



v.a. binomial como suma de v.a. geométricas

Aplicación de las distribuciones binomial negativa y geométrica

- Si la probabilidad de tener que hacer un gran número de intentos antes de obtener un éxito (distrib. geométrica) o r éxitos (distrib. Binomial negativa) es alta, esto podría significar que se emplearían muchos esfuerzos en lograr el objetivo (lo que podría significar altos costos en un proyecto).