60
2EJ

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

PROCESO DE POISSON Y TEOREMAS LIMITE

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A

JUAN CARLOS MEJIA PEREZ

FACULTAD DE CIENCIAS
UNAM

MEXICO, D. F. AGOSTO DE 1995

VNIVERSIDAD NACIONAL
AVPNºMA DE
MEXICO

M. en C. Virginia Abrín Batule
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:

"PROCESO DE POISSON Y TEOREMAS LIMITE"

realizado por JUAN CARLOS MEJIA PEREZ,

con número de cuenta 8707720-3 , pasante de la carrera de ACTUARIA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

DRA. MARIA EMILIA CABALLERO ACOSTA
Director de Tesis
Propietario
DRA. ASUNCION BEGOÑA FERNANDEZ FERNANDEZ
Propietario
MAT. MARGARITA ELVIRA CHAVEZ CANO
Propietario
DR. FEDERICO O'REILLY TONGO
Suplente
M. EN C. RAUL RUEDA DIAZ DEL CAMPO
Suplente

Consejo Departamental de Matemáticas

MAT. CESAR GUEVARA BRAVO

A Dios.

A mis padres, quienes me apoyaron y estuvieron conmigo siempre que los necesité. Quienes me enseñaron a ser independiente.

A mis hermanos: Laura, Patricia, Martha y Mauricio quienes compartieron conmigo tantas y tantas experiencias durante toda su vida.

A mis abuelitos Jesús y Josefina por brindarme lo mejor de su vida.

A mis tios y tias.

A Adriana por todo ese apoyo incondicional que me ha brindado, por su confianza, amor y amistad.

A mis compañeros de la Facultad de Ciencias con quienes compartítantas desveladas y días sin comer.

A María Emilia, por haberme dedicado un poco de su tiempo y por abrirme las puertas de su amistad.

A todos los profesores de la Facultad por brindarme la oportunidad de conocer el maravilloso mundo de las matemáticas, pero muy en especial a Begoña, Angel Carrillo, Agustín y Guadalupe Cano, Margarita Chavez y a Laura Ortiz y Ernesto Rosales.

Contenido

dependiendo de la naturaleza misma del problema. Posteriormente se estudian conceptos tales como tiempos de espera, estadísticas de orden y tiempos entre llegadas. Cabe hacer notar que se puede empezar a leer el capítulo 2 con los temas analizados hasta la sección 1.5. Las secciones restantes se presentan para complementar el estudio realizado acerca del proceso Poisson, no obstante, también se pueden encontrar resultados interesantes que vale la pena revisar.

En el cápitulo 2 se estudia el método de Chen-Stein. Se empieza por dar una visión general del método. Posteriormente se define la notación para poder enunciar el teorema principal, se hacen algunas observaciones importantes y se entra de lleno en la demostración del teorema. Al final de la sección 2.4 se dan un par de lemas que son la base para la demostración del teorema principal. Sus demostraciones son muy extensas, por ello, para no perder la esencia de la demostración del teorema,decidimos que las incluiríamos en un apéndice al final del trabajo. La demostración de estos lemas se puede encontrar en [3], pero debido a su complejidad, se reestructuró para poder hacerla un poco más accesible.

Finalmente, en el capítulo 3 se estudian algunos ejemplos para ayudar al lector a entender mejor cómo se usa el método de Chen-Stein. El primer ejemplo es una aplicación a las gráficas aleatorias. Este ejemplo pudo ser presentado de una manera muy clara gracias a algunas pláticas que tuve con el DR. Victor Hugo de la Peña, de la Universidad de Columbia, Nueva York.En el apéndice 2 se pueden encontrar las definiciones de los conceptos de teoría de gráficas que utilizamos a lo largo del ejemplo.

El segundo ejemplo es un estudio de la aproximación Poisson a la distribución del número de personas dentro de algún grupo que cumplen años el mismo día. Este ejemplo fué ampliado para dar una mayor claridad. Como dato curioso, se intentó recopilar las fechas de cumpleaños de los gobernantes que ha tenido México a lo largo de su historia con el objeto de saber si se obtenía alguna coincidencia en tales fechas de cumpleaños.

En el tercer y último ejemplo presentado en este trabajo se estudia un problema relacionado con las rachas en lanzamientos de monedas (que caigan muchos soles seguidos o muchas águilas).

Introducción

El objetivo principal de este trabajo es presentar los resultados que se encuentran en el artículo titulado *"Two moments suffice for Poisson approximations: The Chen-Stein method"* [2] de una manera sencilla, fácil de entender y, sobre todo, fácil de usar.

La razón principal para estudiar este artículo es que los resultados que ahí se presentan son de gran utilidad cuando podemos expresar nuestra variable aleatoria en términos de una suma de variables aleatorias Bernoulli, no importando si estas son independientes o no. De hecho, cuando las variables aleatorias son dependientes, es cuando el método realmente muestra su potencialidad, pues cuando se tienen variables aleatorias independientes, ya existen teoremas límite que son muy conocidos.

El método que aquí presentamos, no sólo es un teorema límite, sino que además da una cota para el error, lo cual es de gran ayuda en problemas prácticos. Este método al que nos estamos refiriendo es conocido como el *método de Chen-Stein* que, a muy grandes rasgos dice que

> Si uno puede mostrar que "ciertos parámetros" son relativamente "pequeños", entonces se puede utilizar una variable aleatoria Poisson para calcular probabilidades, en lugar de la variable aleatoria con la que originalmente se trabajaba y que tal vez es muy difícil de manejar.

Este trabajo se divide en tres capítulos. En el cápitulo 1 se estudian algunas formas de caracterizar totalmente a una distribución de probabilidad. En la sección 1.3 se dan dos definiciones equivalentes del Proceso Poisson y se demuestra tal equivalencia. Es importante hacer notar que ambas definiciones son muy útiles, cada una

Es importante hacer notar que los ejemplos presentados aquí no son los únicos a los que se les puede aplicar el método de Chen-Stein, de hecho, la gama de problemas es muy amplia. Algunos otros ejemplos pueden ser encontrados en [1] .

Capítulo 1

Preliminares.

1.1 Cómo caracterizar una distribución.

Empezaremos por estudiar algunas formas de caracterizar totalmente una distribución de probabilidad.

I. Función generadora de momentos.

Definición 1.1.1
Sea X una variable aleatoria. La función generadora de momentos de X se define como

$$\Psi_X(t) = E[e^{tX}] = \int e^{tX} dF(X).$$

Donde:

$$\int e^{tX} dF(X) = \begin{cases} \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{si } X \text{ es continua} \\ \\ \sum_x e^{tx} P(X = x) & \text{si } X \text{ es discreta.} \end{cases}$$

Cuando no haya confusión y para efectos de simplicidad, escribiremos $\Psi(t)$ en lugar de $\Psi_X(t)$.

Todos los momentos de X pueden ser obtenidos sucesivamente diferenciando a la función Ψ con respecto a t y evaluando en cero.

Así :

$$\begin{aligned}\Psi'(t) &= E[Xe^{tX}]\\ \Psi''(t) &= E[X^2e^{tX}]\\ &\vdots\\ \Psi^{(n)}(t) &= E[X^ne^{tX}].\end{aligned}$$

Evaluando en $t=0$ tenemos

$$\Psi^{(n)}(0)=E[X^n], \qquad n\geq 1.$$

Nótese que hemos supuesto que podemos intercambiar el operador de integración y el de diferenciación. Este es usualmente el caso.

Cuando una función generadora de momentos existe, ésta determina de manera única a la distribución. Esto es sumamente importante porque nos permite caracterizar la distribución de probabilidad de una variable aleatoria por medio de su función generadora de momentos.

Ejemplos

1. Binomial (n,p).

$$\begin{aligned}\Psi(t) &= \sum_{k=0}^{n} e^{tk}\binom{n}{k}(p)^k(1-p)^{n-k}\\ &= \sum_{k=0}^{n}\binom{n}{k}(pe^t)^k(1-p)^{n-k}\\ &= [pe^t+(1-p)]^n\\ &= [p(e^t-1)+1]^n.\end{aligned}$$

2. Poisson(λ).

$$\begin{aligned}\Psi(t) &= \sum_{k=0}^{\infty} e^{tk}e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!}\\ &= \exp\left\{-\lambda+\lambda e^t\right\} = \exp\left\{\lambda(e^t-1)\right\}.\end{aligned}$$

3. Exponencial(λ).

$$\begin{aligned}\Psi(t) &= \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^\infty e^{(t-\lambda)x}dx \\ &= \frac{\lambda}{t-\lambda}e^{(t-\lambda)x}\Big|_0^\infty = \frac{\lambda}{t-\lambda}.\end{aligned}$$

Así :

$$\Psi(t) = \frac{\lambda}{t-\lambda}, \qquad t<\lambda.$$

4. Normal(0,1).

$$\begin{aligned}\Psi_Z(t) &= \int_{-\infty}^\infty e^{tz}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-(z^2-2tz)/2}dz \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{[-(z-t)^2+t^2]/2}dz \\ &= \frac{e^{t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-y^2/2}dy, \qquad \text{donde } y=z-t. \\ &= e^{t^2/2}.\end{aligned}$$

5. Normal(μ,σ^2).

Sabemos que si Z es una variable aleatoria con distribución normal estándar y X es una variable aleatoria normal con media μ y varianza σ^2, entonces podemos escribir a X como

$$X = \mu + \sigma Z.$$

De esta manera

$$\begin{aligned}E[e^{tX}] &= e^{t\mu}E[e^{t\sigma Z}] = e^{t\mu}\Psi_Z(t\sigma) \\ &= e^{t\mu}e^{(\sigma t)^2/2} = e^{\mu t+(\sigma^2t^2)/2}.\end{aligned}$$

Por lo tanto

$$\Psi_X(t) = e^{\mu t+(\sigma^2t^2)/2}.$$

6. Sean X y Y variables aleatorias independientes normales con medias μ_1,μ_2 y varianzas σ_1^2 y σ_2^2 respectivamente. La función generadora de momentos de la suma

$X+Y$ es

$$\begin{aligned}
\Psi_{X+Y}(t) &= E\left[e^{t(X+Y)}\right] = E\left[e^{tX}\right]E\left[e^{tY}\right] \\
&= \Psi_X(t)\Psi_Y(t) = \exp\left(\mu_1 t + \frac{\sigma_1^2 t^2}{2}\right)\exp\left(\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right) \\
&= \exp\left((\mu_1+\mu_2)t + \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}\right).
\end{aligned}$$

Por lo tanto, $X+Y$ tiene una distribución normal con media $\mu_1+\mu_2$ y varianza $\sigma_1^2+\sigma_2^2$.

7. Notemos que, en general, si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes, la función generadora de momentos de

$$Y = \sum_{i=1}^{n} X_i$$

es

$$\begin{aligned}
\Psi_Y(t) &= E\left[e^{tY}\right] &=& E\left[\exp(t\sum_{i=1}^{n} X_i)\right] \\
&= E\left[\prod_{i=1}^{n} e^{tX_i}\right] &=& \prod_{i=1}^{n} E\left[e^{tX_i}\right] \\
&= \prod_{i=1}^{n} \Psi_{X_i}(t).
\end{aligned}$$

II. Función característica.

Como la función generadora de momentos no siempre existe, es teóricamente conveniente definir otra función alternativa.

Definición 1.1.2
Sea X una variable aleatoria. La función característica de X se define como

$$\phi_X(t) = E\left[e^{itX}\right], \qquad -\infty < t < \infty.$$

Puede ser mostrado que ϕ siempre existe y que, al igual que la función generadora de momentos, determina de manera única la distribución de X.

También podemos definir la función generadora de momentos y la función característica para las variables aleatorias $X_1, \ldots, X_n$ como

$$\begin{aligned} \Psi(t_1,\ldots,t_n) &= E\left[\exp\left(\sum_{j=1}^{n} t_j X_j\right)\right] \\ \phi(t_1,\ldots,t_n) &= E\left[\exp\left(i\sum_{j=1}^{n} t_j X_j\right)\right]. \end{aligned}$$

Puede ser probado que la función generadora de momentos (cuando existe) o la función característica determinan de manera única la función conjunta de probabilidad.

III. Transformada de Laplace.

Cuando se está trabajando con variables aleatorias que sólo toman valores no negativos, a veces es más conveniente utilizar la *transformada de Laplace*, que se define a continuación, en lugar de la función característica.

Definición 1.1.3
Sea X una variable aleatoria con distribución F que toma valores no negativos. La transformada de Laplace para la distribución F se define como

$$\tilde{F}_X(s) = \int_0^\infty e^{-sx} dF(x), \qquad s = a + ib, \quad a \geq 0.$$

Como en el caso de la función característica, la transformada de Laplace determina de manera única la distribución.

1.2 Distribución Poisson.

La distribución de probabilidad Poisson fue introducida por S. D. Poisson en un libro que escribió acerca de las aplicaciones de la teoría de la probabilidad a disputas legales y hechos delictivos. Este libro, publicado en 1837, fue titulado *Recherches sur la probabilité des jugements en matière criminelle et en matière civile.*

Como veremos más adelante, la variable aleatoria Poisson tiene un rango de aplicaciones muy grande en diversas áreas debido a resultados límite en los que esta distribución aparece. Estos resultados se mencionan más adelante

La distribución Poisson con parámetro $\lambda > 0$ está dada por

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0,1,2,\ldots$$

Sea X una variable aleatoria con distribución Poisson y parámetro λ, que denotaremos, en lo sucesivo, como $X \sim Poisson(\lambda)$.

Calculemos E[X] y var(X):

$$\begin{aligned} E[X] &= \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} &=& \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\ &= \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} &=& \lambda e^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} \\ &= \lambda. \end{aligned}$$

Para calcular la varianza, hagamos:

$$E[X^2] = E[X(X-1)] + E[X].$$

Ahora:

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty}\frac{x(x-1)e^{-\lambda}\lambda^x}{x!} &=& \lambda^2 e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} \\ &= \lambda^2 e^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} &=& \lambda^2. \end{aligned}$$

Finalmente

$$var(x) = E[X^2] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Por lo tanto

$$E[X] = var(X) = \lambda. \tag{1}$$

A continuación enunciaremos un teorema que relaciona de manera muy interesante la distribución de una suma de variables aleatorias Poisson independientes con su parámetro.

TEOREMA 1.2.1
Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con distribución Poisson y parámetros $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectivamente. Entonces

$$\sum_{i=1}^{n} X_i \sim Poisson\left(\sum_{i=1}^{n} \lambda_i\right).$$

Demostración:
(La demostración se hará por inducción)
Para $n = 2$:

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{s=0}^{k} P(X_1 = s, X_2 = k - s) \\
&= \sum_{s=0}^{k} P(X_1 = s)P(X_2 = k - s) \\
&= \sum_{s=0}^{k} \left(\frac{e^{-\lambda_1}\lambda_1^s}{s!} \cdot \frac{e^{-\lambda_2}\lambda_2^{k-s}}{(k-s)!}\right) \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{s=0}^{k} \frac{k!}{s!(k-s)!} \lambda_1^s \lambda_2^{k-s} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1 + \lambda_2)^k}{k!}.
\end{aligned}$$

Donde la última igualdad se sigue del teorema del binomio.
Supongamos que es válido para $n = k$. Demostraremos que es válido para $n = k + 1$. $\sum_{i=1}^{k+1} X_i = \sum_{i=1}^{k} X_i + X_{k+1}$, pero $\sum_{i=1}^{k} X_i \sim$ Poisson($\sum_{i=1}^{k} \lambda_i$) por hipótesis de inducción, y $X_{k+1} \sim$ Poisson($\lambda_{k=1}$). Por lo tanto, $\sum_{i=1}^{k+1} X_i \sim$ Poisson($\sum_{i=1}^{k+1} \lambda_i$).

□

Consideremos una variable aleatoria Poisson N con parámetro $\mu > 0$ y escribamos a N como la suma de unos:

$$N = 1 + 1 + \cdots + 1.$$

Supongamos que cada sumando toma el valor 1 con probabilidad p y 0 con probabilidad $1-p$. ¿Cuál es la Distribución de la suma M si ésta es de la forma $M = 1+0+0+\cdots+1$? El siguiente teorema establece la respuesta de manera precisa.

TEOREMA 1.2.2
Sean M, N variables aleatorias tales que $N \sim$Poisson(λ) y $(M \mid N)$ tiene una distribución Binomial(n,p), entonces $M \sim$Poisson(λp).

Demostración:

$$\begin{aligned}
P(M=k) &= \sum_{n=0}^{\infty} P(M=k, N=n) &&= \sum_{n=0}^{\infty} P(M=k \mid N=n)P(N=n) \\
&= \sum_{n=0}^{\infty}\left(\frac{n!p^k(1-p)^{n-k}}{k!(n-k)!}\cdot\frac{e^{-\lambda}\lambda^n}{n!}\right) &&= \frac{e^{-\lambda}(p\lambda)^k}{k!}\sum_{n=k}^{\infty}\frac{[\lambda(1-p)]^{n-k}}{(n-k)!} \\
&= \frac{e^{-\lambda}(p\lambda)^k}{k!}\sum_{m=0}^{\infty}\frac{[\lambda(1-p)]^m}{m!} &&= \frac{e^{-\lambda p}(\lambda p)^k}{k!}.
\end{aligned}$$

Por lo tanto $M \sim$ Poisson(λp).

□

1.3 Proceso Poisson.

En esta sección daremos un par de definiciones de proceso Poisson y probaremos su equivalencia. Para ello introduciremos el concepto de proceso estocástico y proceso de conteo.

Definición 1.3.1
Un proceso estocástico $\underline{X} = \{X(t), t \in T\}$ es una colección de variables aleatorias. Esto es, para cada t en el conjunto de índices $T, X(t)$ es una variable aleatoria.

A menudo interpretamos a t como el tiempo y decimos que $X(t)$ es el estado al tiempo t.

Si el conjunto de índices T es contable, $\underline{X}$ es llamado un proceso estocástico con tiempo discreto; si T es continuo, $\underline{X}$ es llamado proceso estocástico con tiempo continuo.

Definición 1.3.2
Se dice que un proceso estocástico $\underline{X} = \{X(t), t \in T\}$ con tiempo continuo tiene incrementos independientes si, para todo $t_0, t_1, \ldots, t_n$, las variables aleatorias

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$

son independientes.

Se dice que el proceso tiene incrementos estacionarios si $X(t+s) - X(t)$ tiene la misma distribución para todo $t \in T$.

Definición 1.3.3
Sea $\{N(t), t \geq 0\}$ un proceso estocástico. Si $N(t)$ representa el número total de eventos al tiempo t, $\{N(t), t \geq 0\}$ es llamado un proceso de conteo.

Así , un proceso de conteo debe satisfacer:
$i)$ $N(t) \geq 0$.
$ii)$ $N(t)$ toma valores enteros.
$iii)$ Si $s < t \Longrightarrow N(s) \leq N(t)$.
$iv)$ Para $s < t, N(t) - N(s)$ representa el número total de eventos en el intervalo $(s, t]$.
Ahora si estamos listos para definir al proceso Poisson.

Definición 1.3.4
El proceso de conteo $\{N(t), t \geq 0\}$ es llamado proceso Poisson con tasa $\lambda > 0$ si:
$i)$ $N(0) = 0$.
$ii)$ *El proceso tiene incrementos independientes.*
$iii)$ *El número de eventos en cualquier intervalo de longitud t se distribuye Poisson con tasa λt. Esto es, para todo $s, t \geq 0$ se cumple que*

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

Notemos que se sigue de la condición $iii)$ de la definición anterior y de la condición (1) de la página 6 que un proceso Poisson tiene incrementos estacionarios y que $E[N(t)] = \lambda t$.

Antes de dar la siguiente definición de Proceso Poisson necesitamos introducir el siguiente concepto:

Definición 1.3.5
Decimos que la función f es una función $o(h)$ si

$$\lim_{h\to 0}\frac{f(h)}{h}=0.$$

Ahora sí estamos listos para dar una definición equivalente.

Definición 1.3.6
El proceso de conteo $\{N(t), t\geq 0\}$ es llamado proceso de Poisson con tasa $\lambda > 0$ si:
$i)N(0)=0$.
$ii)$ El proceso tiene incrementos estacionarios e independientes.
$iii)P(N(h)=1)=\lambda h+o(h)$.
$iv)P(N(h)\geq 2)=o(h)$.

TEOREMA 1.3.1
Las definiciones 1.3.4 y 1.3.6 son equivalentes.

Demostración:
Primero demostraremos que la definición 1.3.6 implica la definición 1.3.4 :
Sea $P_n(t)=P(N(t)=n)$.
Obtengamos una ecuación diferencial para $P_0(t)$ de la siguiente manera:

$$\begin{aligned}
P_0(t+h) &= P(N(t+h)=0)\\
&= P(N(t)=0, N(t+h)-N(t)=0)\\
&= P(N(t)-N(0)=0, N(t+h)-N(t)=0)\\
&= P(N(t)-N(0)=0)P(N(t+h)-N(t)=0)\\
&= P(N(t)=0)P(N(t+h)-N(t)=0)\\
&= P_0(t)P(N(h)-N(0)=0)\\
&= P_0(t)P(N(h)=0).
\end{aligned}$$

Ahora

$$\begin{aligned}
P(N(h)=0) &= 1-[P(N(h)=1)+P(N(h)\geq 2)]\\
&= 1-[\lambda h+o(h)+o(h)]\\
&= 1-[\lambda h+o(h)]=1-\lambda h+o(h).
\end{aligned}$$

Por lo tanto

$$P_0(t+h) = P_0(t)[1 - \lambda h + o(h)] = P_0(t) - \lambda h P_0(t) + o(h).$$

Así, tenemos que

$$P_0(t+h) - P_0(t) = -\lambda h P_0(t) + o(h).$$

Dividiendo por h y haciendo $h \to 0$, obtenemos la ecuación diferencial

$$P_0'(t) = -\lambda P_0(t)$$

con condición inicial $P_0(0) = 1$, lo cual nos da como solución

$$P_0(t) = e^{-\lambda t}.$$

Similarmente, cuando $n \geq 1$

$$\begin{aligned}
P_n(t+h) &= P(N(t+h) = n) \\
&= P(N(t) = n, N(t+h) - N(t) = 0) \\
&+ P(N(t) = n-1, N(t+h) - N(t) = 1) \\
&+ P(N(t+h) - N(t) \geq 2, N(t+h) = n) \\
&= P(N(t) = n)P(N(t+h) - N(t) = 0) \\
&+ P(N(t) = n-1)P(N(t+h) - N(t) = 1) + o(h) \\
&= P_n(t)[1 - \lambda h + o(h)] + P_{n-1}(t)[\lambda h + o(h)] + o(h) \\
&= P_n(t) - \lambda h P_n(t) + \lambda h P_{n-1}(t) + o(h).
\end{aligned}$$

Por lo tanto

$$P_n(t+h) - P_n(t) = -\lambda h P_n(t) + \lambda h P_{n-1}(t) + o(h).$$

Dividiendo por h y haciendo $h \to 0$ tenemos que

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t).$$

De aquí que

$$e^{\lambda t}[P_n'(t) + \lambda P_n(t)] = \lambda e^{\lambda t} P_{n-1}(t),$$

y, finalmente

$$\frac{d}{dt}\left(e^{\lambda t}P_n(t)\right) = \lambda e^{\lambda t}P_{n-1}(t).$$

Haremos inducción sobre n.

Cuando $n = 1$

$$\frac{d}{dt}\left(e^{\lambda t}P_1(t)\right) = e^{\lambda t}P_0(t) = \lambda.$$

Y al resolver la ecuación diferencial con condición inicial $P_1(0) = 0$ obtenemos que

$$P_1(t) = \lambda t e^{-\lambda t}.$$

Para terminar el proceso de inducción, supongamos válido para $n = k$, es decir,

$$P_k(t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

Vamos a demostrar que es válido para $n = k + 1$

$$\begin{aligned}\frac{d}{dt}\left(e^{\lambda t}P_{k+1}(t)\right) &= \lambda e^{\lambda t}P_k(t)\\ &= \frac{\lambda(\lambda t)^k}{k!}\end{aligned}$$

Finalmente, resolviendo la ecuación con condición inicial $P_{k+1}(t) = 0$ llegamos a la ecuación

$$P_{k+1}(t) = \frac{e^{-\lambda t}(\lambda t)^{k+1}}{(k+1)!}.$$

Con lo cual concluimos la inducción. Por lo tanto

$$P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}. \qquad \forall n \in \mathbb{N}$$

Pero $P(N(t) = n) = P(N(t) - N(0) = n) = P(N(t + h) - N(h) = n)$ para todo número real h no negativo (Por el inciso ii)). Por lo tanto

$$P(N(t + s) - N(s) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \qquad \forall s, t \geq 0.$$

Así, se cumple el inciso iii) de la definición 1.3.4, con lo cual queda demostrada la primera parte del teorema.

Ahora demostraremos que la definición 1.3.4 implica la definición 1.3.6.

Basta probar los incisos (*iii*) y (*iv*).
inciso (iii):

$$\begin{aligned} P(N(h)=1) &= P(N(h)-N(0)=1)=\lambda h e^{-\lambda h} \\ &= \lambda h \sum_{i=0}^{\infty} \frac{(-\lambda h)^i}{i!} \qquad \text{(Por expansión de Taylor.)} \\ &= \lambda h[1-\lambda h+\frac{(\lambda h)^2}{2!}-\cdots]=\lambda h+\lambda h[-\lambda h+\frac{(\lambda h)^2}{2!}-\cdots]. \end{aligned}$$

Pero $\lambda h\left(-\lambda h+\frac{(\lambda h)^2}{2!}-\cdots\right)=o(h)$.
Así finalmente $P(N(h)=1)=\lambda h+o(h)$.
inciso (iv):

$$\begin{aligned} P(N(h)\geq 2) &= 1-P(N(h)=0)-P(N(h)=1)=1-e^{-\lambda h}-\lambda h+o(h) \\ &= 1-[1-\lambda h+o(h)]-\lambda h+o(h)=o(h) \end{aligned}$$

Con lo cual queda completa la demostración del teorema.

□

1.4 Tiempos entre llegadas.

Definición 1.4.1
Consideremos un proceso Poisson, y sea X_1 el momento en el que ocurre el primer evento. Además, para $n \geq 1$, X_n denota el tiempo que transcurre entre el $(n-1)$-ésimo y el n-ésimo evento. La sucesión $\{X_n, n \geq 1\}$ es llamada sucesión de tiempos entre llegadas.

Para determinar la distribución de X_n primero notemos que el evento $\{X_1 > t\}$ ocurre si y sólo si no hay eventos en el intervalo $[0,t]$, así :

$$P(X_1 > t) = P(N(t)=0) = e^{-\lambda t}.$$

Lo cual indica que X_1 tiene una distribución exponencial con media $1/\lambda$.

Ahora calculemos la distribución conjunta de X_1 y X_2. Primero notemos que

$$\begin{aligned}
P(X_1 \leq t_1, X_2 \leq t_2) &= E\left[1_{\{X_1 \leq t_1, X_2 \leq t_2\}}\right] \\
&= E\left[E\left[1_{\{X_1 \leq t_1, X_2 \leq t_2\}} | X_1\right]\right] \\
&= \int_0^{\infty} E\left[1_{\{X_1 \leq t_1, X_2 \leq t_2\}} | X_1 = s\right] \lambda e^{-\lambda s} ds \\
&= \int_0^{\infty} P(X_1 \leq t_1, X_2 \leq t_2 | X_1 = s) \lambda e^{-\lambda s} ds
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
P(X_1 \leq t_1, X_2 \leq t_2) &= \int_0^{\infty} P(X_1 \leq t_1, X_2 \leq t_2 | X_1 = s) \lambda e^{-\lambda s} ds \\
&= \int_0^{\infty} P(s \leq t_1, X_2 \leq t_2 | X_1 = s) \lambda e^{-\lambda s} ds \\
&= \int_0^{t_1} P(X_2 \leq t_2 | X_1 = s) \lambda e^{-\lambda s} ds \\
&= \int_0^{t_1} [1 - P(X_2 \geq t_2 | X_1 = s)] \lambda e^{-\lambda s} ds \\
&= \int_0^{t_1} [1 - P(N(s + t_2) - N(s) = 0)] \lambda e^{-\lambda s} ds \\
&= \int_0^{t_1} \left(1 - e^{-\lambda t_2}\right) \lambda e^{-\lambda s} ds \\
&= \left(1 - e^{-\lambda t_2}\right)\left(1 - e^{-\lambda t_1}\right)
\end{aligned}$$

Lo cuál quiere decir que X_1 y X_2 son variables aleatorias independientes y que tienen distribución exponencial con media $1/\lambda$. Estas consideraciones nos motivan a enunciar la siguiente

Proposición 1.4.1

Sea X_n la variable que mide el tiempo entre llegada del $(n-1)$-ésimo evento y el n-ésimo, para $n = 1, 2, \ldots$, entonces $X_1, X_2, \ldots$, son variables aleatorias independientes con distribución exponencial y media $1/\lambda$.

Demostración:

(La demostración ser hará por inducción.)

Para $n = 2$ se demostró arriba.

Supongamos que el resultado es válido para $n-1$, es decir, $X_1,\ldots,X_{n-1}$ son variables aleatorias independientes con distribución exponencial y media $1/\lambda$. Demostraremos que es también válido para n.

$$
\begin{aligned}
& P(X_1 \le t_1,\ldots,X_n \le t_n) \\
= & \int_0^{\infty}\cdots\int_0^{\infty} P(X_1 \le t_1,\ldots,X_n \le t_n | X_1 = s_1,\ldots,X_{n-1} = s_{n-1}) \\
\times & \lambda^{n-1} e^{-\lambda(s_1+\cdots+s_{n-1})} ds_1\cdots ds_{n-1} \\
= & \int_0^{\infty}\cdots\int_0^{\infty} P(s_1 \le t_1,\cdots,s_{n-1} \le t_{n-1} | X_1 = s_1,\ldots,X_{n-1} = s_{n-1}) \\
\times & \lambda^{n-1} e^{-\lambda(s_1+\cdots+s_{n-1})} ds_1\cdots ds_{n-1} \\
= & \int_0^{t_{n-1}}\cdots\int_0^{t_1} P(X_n \le t_n | X_1 = s_1,\ldots,X_{n-1} = s_{n-1}) \\
\times & \lambda^{n-1} e^{-\lambda(s_1+\cdots+s_{n-1})} ds_1\cdots ds_{n-1} \\
= & \int_0^{t_{n-1}}\cdots\int_0^{t_1} [1 - P(X_n \ge t_n | X_1 = s_1,\ldots,X_{n-1} = s_{n-1})] \\
\times & \lambda^{n-1} e^{-\lambda(s_1+\cdots+s_{n-1})} ds_1\cdots ds_{n-1} \\
= & \int_0^{t_{n-1}}\cdots\int_0^{t_1} \left(1 - e^{-\lambda t_n}\right) \lambda^{n-1} e^{-\lambda(s_1+\cdots+s_{n-1})} ds_1\cdots ds_{n-1} \\
= & \prod_{i=1}^{n} \left(1 - e^{-\lambda t_i}\right)
\end{aligned}
$$

Por consiguiente $X_1,\ldots,X_n$ son variables aleatorias independientes exponenciales con media $1/\lambda$.

□

Observación 1.4.1

La suposición que se hace acerca de los incrementos independientes y estacionarios del proceso nos aseguran que, en cualquier punto del tiempo, el proceso vuelve a comenzar y tiene la misma distribución que el proceso original. En otras palabras, el proceso no tiene memoria, por ello, era de esperarse que apareciera la distribución exponencial.

Otra cantidad interesante es S_n, el tiempo en el que ocurre el n-ésimo evento, también llamada tiempo de espera hasta el n-ésimo evento, así

$$S_n = \sum_{i=1}^{n} X_i, \qquad n \ge 1.$$

Proposición 1.4.2

$$S_n \sim Gama(n, \lambda).$$

Demostración:
Utilizaremos la definición 1.1.1 de la página 1.
La función generadora de momentos de X_i es $\lambda/(\lambda - t), t < \lambda, \forall i = 1, 2, \ldots, n$.
Calculemos la función generadora de momentos de S_n

$$\begin{aligned} E\left[e^{tS_n}\right] &= E\left[\exp\{t\sum_{i=1}^{n} X_i\}\right] = E\left[\exp\{\sum_{i=1}^{n} tX_i\}\right] \\ &= E\left[\prod_{i=1}^{n} e^{tX_i}\right] \end{aligned}$$

pero $\{X_n\}$ es una sucesión de variables aleatorias independientes, por lo tanto

$$\begin{aligned} E\left[e^{tS_n}\right] &= E\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} E\left[e^{tX_i}\right] \\ &= \left(\frac{\lambda}{\lambda - t}\right)^n, \end{aligned}$$

que es la función generadora de momentos de una variable aleatoria $Gama(n, \lambda)$. Esto es, la función de distribución es

$$f(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t} \quad = \quad \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \tag{2}$$

pues como $n \in \mathbb{N}$ entonces $\Gamma(n) = (n-1)!$.

□

Otra manera de hacer la demostración es la siguiente:

$$F_{S_n}(t) = P(S_n \leq t) = P(N(t) \geq n) = \sum_{j=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^j}{j!}$$

por lo tanto

$$f_{S_n}(t) \quad = \quad F'_{S_n}(t) = \sum_{j=n}^{\infty} \frac{1}{j!} \left[\lambda j e^{-\lambda t}(\lambda t)^{j-1} - \lambda e^{-\lambda t}(\lambda t)^j\right]$$

$$\begin{aligned}
&= \sum_{j=n}^{\infty}\left[\lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!} - \lambda e^{-\lambda t}\frac{(\lambda t)^{j}}{j!}\right] \\
&= \lambda e^{-\lambda t}\left[\sum_{j=n}^{\infty}\frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty}\frac{(\lambda t)^{j}}{j!}\right] \\
&= \lambda e^{-\lambda t}\left[\frac{(\lambda t)^{n-1}}{(n-1)!}\right].
\end{aligned}$$

Por lo tanto

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

que es la misma que la obtenida en la ecuación (2) .

De acuerdo a las consideraciones anteriores, podemos dar otra definición de proceso Poisson de la siguiente manera: Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias exponenciales independientes exponenciales con media $1/\lambda$. Definimos un proceso de conteo diciendo que el n-ésimo evento del proceso ocurre al tiempo S_n, donde

$$S_n = X_1 + X_2 + \cdots + X_n.$$

El proceso de conteo resultante será un proceso Poisson con tasa λ.

1.5 Tiempos de espera y estadísticas de orden.

Supongamos que sabemos que un evento de un proceso Poisson ha tenido lugar al tiempo t; queremos conocer la distribución del tiempo al cual ocurre tal evento. Dado que el proceso Poisson posee incrementos estacionarios e independientes, parece razonable que cualquier intervalo en $[0, t]$ de la misma longitud debería tener la misma probabilidad de contener al evento. En otras palabras, el tiempo del evento debería distribuirse uniformemente sobre el intervalo $[0, t]$. Esto se puede checar fácilmente dado que, para $s \leq t$,

$$\begin{aligned}
P(X_1 < s \mid N(t) = 1) &= \frac{P(\text{Un evento en } [0, s) \text{ y ninguno en } [s, t])}{P(N(t) = 1)} \\
&= \frac{P(\text{Un evento en } [0, s))P(\text{No haya evento en } [s, t])}{P(N(t) = 1)}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\left(\lambda s e^{-\lambda s}\right)\left(e^{-\lambda(t-s)}\right)}{\lambda t e^{-\lambda t}} \\
&= \frac{s e^{-\lambda t}}{t e^{-\lambda t}} \\
&= \frac{s}{t}.
\end{aligned}$$

Este resultado puede ser generalizado pero, para ello, necesitamos introducir el concepto de ***estadísticas de orden***.

Definición 1.5.1
Sean $Y_1, \ldots, Y_n$, n variables aleatorias con distribución común F. Definamos

$$\begin{aligned}
Y_{(1)} &= \textit{El valor más pequeño de } \{Y_1, \ldots, Y_n\}. \\
Y_{(2)} &= \textit{El segundo valor más pequeño de } \{Y_1, \ldots, Y_n\} \\
&\vdots \\
Y_{(j)} &= \textit{El } j-\textit{ésimo valor más pequeño de } \{Y_1, \ldots, Y_n\} \\
&\vdots \\
Y_{(n)} &= \textit{El valor más grande de } \{Y_1, \ldots, Y_n\}.
\end{aligned}$$

Los valores ordenados $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ son conocidos como las estadísticas de orden correspondientes a las variables aleatorias $Y_1, \ldots, Y_n$. En otras palabras, $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ son los valores ordenados de manera creciente de $\{Y_1, \ldots, Y_n\}$.

Si las Y_i's son variables aleatorias independientes e idénticamente distribuidas con densidad de probabilidad f, entonces la densidad conjunta de las estadísticas de orden $Y_{(1)}, \ldots, Y_{(n)}$ está dada por

$$f(y_1, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i), \qquad y_1 < \cdots < y_n. \tag{3}$$

Lo anterior se debe a que
$i)$ $(Y_1, \ldots, Y_n)$ será igual a $(y_1, \ldots, y_n)$ si $(Y_1, \ldots, Y_n)$ es igual a cualquiera de las $n!$ permutaciones de $\{y_1, \ldots, y_n\}$.

ii) La densidad de probabilidad de $f_{Y_1 \cdots Y_n}(y_1, \ldots, y_n)$ es igual a

$$f_{Y_1}(y_1) \cdots f_{Y_n}(y_n) = \prod_{i=1}^{n} f_{Y_i}(y_i).$$

Si las Y_i's,$i = 1, \ldots, n$, son uniformemente distribuidas sobre el intervalo $(0, t)$, se sigue de la ecuación (3) que la distribución conjunta de las estadísticas de orden $(Y_{(1)}, \ldots Y_{(n)})$ es

$$f(y_1, \ldots, y_n) = \frac{n!}{t^n}.$$

Ahora estamos listos para enunciar el siguiente teorema, el cual tiene aplicaciones muy interesantes.

TEOREMA 1.5.1

Dado que $N(t) = n$, los tiempos de espera $S_1, \ldots, S_n$ tienen la misma distribución que las estadísticas de orden correspondientes a n variables aleatorias uniformemente distribuidas en el intervalo $(0, t)$.

Demostración:

Sean $0 < t_1 < t_2 < \cdots < t_{n+1} = t$ y sea h_i suficientemente pequeño para que $t_i + h_i < t_{i+1}, i = 1, \ldots, n$. Ahora, si representamos por A al evento

> Hay exactamente un evento en el intervalo $[t_i, t_i + h_i]$, para $i = 1, \ldots, n$ y no haya más eventos en $[0, t]$,

entonces

$$\begin{aligned} P(t_i \leq S_i \leq t_i + h_i, i = 1, \ldots, n \quad | \quad N(t) = n) &= \frac{P(A)}{P(N(t) = n)} \\ &= \frac{\lambda h_1 e^{-\lambda h_1} \cdots \lambda h_n e^{-\lambda h_n} e^{-\lambda(t - h_1 - h_2 - \cdots - h_n)}}{\frac{(\lambda t)^n}{n!} e^{-\lambda t}} \\ &= \frac{n!}{t^n} h_1 \cdot h_2 \cdots h_n. \end{aligned}$$

Por lo tanto

$$\frac{P(t_i \leq S_i \leq t_i + h_i, i = 1, \ldots, n \,|\, N(t) = n)}{h_1 \cdot h_2 \cdots h_n} = \frac{n!}{t^n}.$$

Y haciendo $h_i \to 0, i = 1, \ldots, n$, tenemos que

$$f_{S_1 \cdots S_n}(t_1, \ldots, t_n) = \frac{n!}{t^n} \qquad 0 < t_1 < \cdots < t_n.$$

□

Nota: Intuitivamente decimos que, bajo la condición de que n eventos han ocurrido en el intervalo $(0,t)$, los tiempos $S_1, \ldots, S_n$ en los cuales ocurren los eventos, considerados como variables aleatorias desordenadas, se distribuyen independiente y uniformemente en el intervalo $(0,t)$.

Ejemplo 1.5.1

Supongamos que los clientes que llegan a una estación de tren lo hacen de acuerdo a un proceso de Poisson con tasa λ. Si el tren parte al tiempo t, entonces la suma esperada de tiempos de espera de los clientes que llegan a la estación en el intervalo $(0,t)$ es:

$$E\left[\sum_{n=1}^{N(t)}(t-S_n)\right] = E\left[E\left[\sum_{n=1}^{N(t)}(t-S_n)\middle| N(t)\right]\right].$$

Ahora calculemos la esperanza condicional utilizando el teorema 1.5.1 :

$$\begin{aligned}
E\left[\sum_{n=1}^{N(t)}(t-S_n)\middle| N(t)=m\right] &= E\left[\sum_{n=1}^{m}(t-S_n)\middle| N(t)=m\right] \\
&= E\left[mt-\sum_{n=1}^{m}S_n\middle| N(t)=m\right] \\
&= mt-E\left[\sum_{n=1}^{m}S_n\middle| N(t)=m\right] \\
&= mt-E\left[\sum_{n=1}^{m}U_{(n)}\right] = mt-E\left[\sum_{n=1}^{m}U_n\right] \\
&= mt-\sum_{n=1}^{m}E[U_n] = mt-m\frac{t}{2} \\
&= \frac{mt}{2}.
\end{aligned}$$

Donde U_i es una variable aleatoria uniformemente distribuida en el intervalo $(0,t)$, $i = 1, \ldots, n$.

Así

$$E\left[\sum_{n=1}^{N(t)}(t-S_n)\middle| N(t)\right] = \frac{tN(t)}{2},$$

por lo tanto

$$E\left[E\left[\sum_{n=1}^{N(t)}(t-S_n)\middle| N(t)\right]\right] = \frac{t}{2}E[N(t)] = \frac{t}{2}\lambda t.$$

Finalmente

$$E\left[\sum_{n=1}^{N(t)}(t-S_n)\right] = \frac{\lambda t^2}{2}.$$

Como una aplicación importante del teorema 1.5.1 supongamos que cada evento de un Proceso Poisson con tasa λ se clasifica en uno de dos tipos de evento: *Tipo I y Tipo II*, dependiendo del momento en el que ocurre. Específicamente, supongamos que si un evento ocurre al tiempo s entonces, independientemente de todo lo demás, se clasifica como evento tipo I con una probabilidad $P(s)$ como tipo II con probabilidad $1-P(s)$. Usando el teorema 1.5.1 podemos probar la siguiente

Proposición 1.5.1 *Si $N_i(t)$ representa el número de eventos tipo i, $i=1,2$ que han ocurrido al tiempo t, entonces $N_1(t)$ y $N_2(t)$son variables aleatorias independientes Poisson con media λtp y $\lambda t(1-p)$ respectivamente, donde*

$$p = \frac{1}{t}\int_0^t P(s)ds.$$

Demostración:

Calculemos la densidad conjunta:

$$\begin{aligned} P(N_1(t)=n, N_2(t) &= m) = \sum_{k=0}^{\infty} P(N_1(t)=n, N_2(t)=m \mid N(t)=k)P(N(t)=k) \\ &= P(N_1(t)=n, N_2(t)=m \mid N(t)=m+n)P(N(t)=m+n). \end{aligned}$$

pues el resto de los términos de la serie es cero.

Ahora consideremos un evento arbitrario en el intervalo $(0,t)$. Por el teorema 1.5.1 este evento ocurrirá en un tiempo uniformemente distribuido en el intervalo $(0,t)$.

Supongamos que el evento ocurre al tiempo s. Debido a las condiciones anteriores, el evento se clasifica como tipo I con una probabilidad

$$p = \frac{1}{t}\int_0^t P(s)ds$$

independientemente de cualquier otro evento. Así,

$$P(N_1(t) = n, N_2(t) = m \mid N(t) = m + n)$$

se puede pensar como n éxitos y m fracasos en $n + m$ ensayos, es decir,

$$P(N_1(t) = n, N_2(t) = m \mid N(t) = m + n) = \binom{n+m}{n} p^n (1-p)^m.$$

Finalmente tenemos que

$$P(N_1(t) = n, N_2(t) = m) = \left[\frac{(\lambda tp)^n}{n!}e^{-\lambda tp}\right]\left[\frac{(\lambda t(1-p))^n}{n!}e^{-\lambda t(1-p)}\right].$$

□

La importancia de la proprosición anterior se ilustra con el siguiente

Ejemplo 1.5.2

LA COLA POISSON CON INFINITOS SERVIDORES.

Supongamos que los clientes llegan a una estación de servicio de acuerdo a un proceso Poisson con tasa λ. A su llegada, los clientes son servidos inmediatamente por uno de un número infinito de servidores y supongamos que los tiempos de servicio son independientes con una distribución común G.

Para calcular la distribución conjunta del número de clientes que han completado su servicio al tiempo t y del número de aquellos que aún no lo completan, diremos que un cliente es del tipo-I si ya ha completado su servicio y del tipo-II en otro caso. Ahora, si el cliente llega al tiempo $s, s \leq t$, entonces será del tipo-I si termina su servicio en un tiempo menor que $t - s$, y dado que la distribución del tiempo de servicio es G, la probabilidad será $G(t - s)$. Así :

$$P(s) = G(t-s), \qquad s \leq t.$$

Sean $N_1(t)$ el número de clientes que han completado su servicio al tiempo t y $N_2(t)$ el número de clientes que no han sido servidos al tiempo t.

Por la proposición anterior obtenemos que la distribución del número de clientes que han completado su servicio al tiempo t, es Poisson con media

$$E[N_1(t)] = \lambda \int_0^t G(t-s)ds = \lambda \int_0^t G(y)dy.$$

Análogamente el número de clientes que no han sido servidos al tiempo t tiene una distribución Poisson con media

$$E[N_2(t)] = \lambda \int_0^t \bar{G}(y)dy.$$

Donde
$\bar{G}(y) = 1 - G(y)$.

Además $N_1(t)$ y $N_2(t)$ son independientes.

Ahora analicemos otro ejemplo.

Ejemplo 1.5.3 *Supongamos que un aparato está sujeto a shocks que ocurren de acuerdo a un proceso Poisson con tasa λ. El i-ésimo shock provoca un daño D_i. Se supone que las D_i's son variables aleatorias independientes e idénticamente distribuidas y también independientes de $\{N(t), t \geq 0\}$, donde $N(t)$ denota el número de shocks en el intervalo $[0,t]$. Se supone que el daño debido al shock decrece exponencialmente en el tiempo, esto es, si el shock produce un daño inicial D, entonces al tiempo t despues de tal daño es $De^{-\alpha t}$ para alguna $\alpha > 0$.*

Si suponemos que el daño es aditivo, entonces $D(t)$, el daño al tiempo t puede ser expresado como:

$$D(t) = \sum_{i=1}^{N(t)} D_i e^{-\alpha(t-S_i)}.$$

Donde S_i es el tiempo en el que ocurre el i-ésimo daño. Podemos determinar $E[D(t)]$ como:

$$E[D(t)] = E\left[E\left[\sum_{i=1}^{N(t)} D_i e^{-\alpha(t-S_i)} \middle| N(t)\right]\right].$$

Ahora

$$\begin{aligned}
E[D(t) \mid N(t)=n] &= E\left[\sum_{i=1}^{N(t)} D_i e^{-\alpha(t-S_i)} \middle| N(t)=n\right] \\
&= E\left[\sum_{i=1}^{n} D_i e^{-\alpha(t-S_i)} \middle| N(t)=n\right] \\
&= \sum_{i=1}^{n} E\left[D_i e^{-\alpha(t-S_i)} \middle| N(t)=n\right] \\
&= \sum_{i=1}^{n} E[D_i \mid N(t)=n] E\left[e^{-\alpha(t-S_i)} \mid N(t)=n\right].
\end{aligned}$$

Pero D_i es independiente del proceso para $i = 1, 2, \ldots, n$ y son variables independientes e idénticamente distribuidas, por lo tanto

$$E[D_i \mid N(t)=n] = E[D],$$

donde D es una variable aleatoria con la misma distribución que las D_i's. Así

$$\begin{aligned}
\sum_{i=1}^{n} E[D_i \mid N(t)=n] E\left[e^{-\alpha(t-S_i)} \mid N(t)=n\right] &= E[D] \sum_{i=1}^{n} E\left[e^{-\alpha(t-S_i)} \mid N(t)=n\right] \\
&= E[D] e^{-\alpha t} \sum_{i=1}^{n} E\left[e^{\alpha S_i} \mid N(t)=n\right] \\
&= E[D] e^{-\alpha t} E\left[\sum_{i=1}^{n} e^{\alpha S_i} \middle| N(t)=n\right].
\end{aligned}$$

Sean $U_1, \ldots, U_n$ variables aleatorias independientes uniformemente distribuidas sobre el intervalo $[0,t]$. Por el teorema 1.5.1 :

$$E\left[\sum_{i=1}^{n} e^{\alpha S_i} \middle| N(t)=n\right] = E\left[\sum_{i=1}^{n} e^{\alpha U_{(i)}}\right].$$

Por lo tanto

$$E[D] e^{-\alpha t} E\left[\sum_{i=1}^{n} e^{\alpha S_i} \middle| N(t)=n\right] = E[D] e^{-\alpha t} E\left[\sum_{i=1}^{n} e^{\alpha U_{(i)}}\right].$$

Ahora:

$$\begin{aligned} E\left[\sum_{i=1}^{n} e^{\alpha U_{(i)}}\right] &= E\left[\sum_{i=1}^{n} e^{\alpha U_i}\right] = nE\left[e^{\alpha U_i}\right] \\ &= \frac{n}{t}\int_0^t e^{\alpha u}du = \frac{n}{\alpha t}e^{\alpha u}\Big|_0^t \\ &= \frac{n}{\alpha t}\left[e^{\alpha t}-1\right] \end{aligned}$$

Así

$$E\left[\sum_{i=1}^{n} e^{\alpha S_{i)}}\middle| N(t)=n\right] = \frac{n}{\alpha t}\left[e^{\alpha t}-1\right].$$

Por lo tanto:

$$E[D(t)\mid N(t)=n] = E[D]e^{-\alpha t}\frac{n}{\alpha t}\left[e^{\alpha t}-1\right] = E[D]\frac{n}{\alpha t}\left[1-e^{-\alpha t}\right].$$

Lo cual nos lleva a

$$E[D(t)\mid N(t)] = E[D]\frac{N(t)}{\alpha t}\left[1-e^{-\alpha t}\right].$$

Finalmente, al tomar esperanza de ambos lados en la ecuación anterior tenemos que

$$E[D(t)] = \frac{\lambda E[D]}{\alpha}\left(1-e^{-\alpha t}\right).$$

Nota: Otra aproximación para obtener $E[D(t)]$ es partir el intervalo $(0,t)$ en intervalos de longitud h que no se intersecten y que lo cubran y entonces sumar la contribución al tiempo t de los shocks que ocurrieron en ese intervalo. Más específicamente, sea h dado y sea X_i la suma de los daños ocurridos al tiempo t que llegaron en el intervalo $I_j=(jh,(j+1)h)$, $\quad j=0,1,\ldots,[t/h]$[1].
De lo anterior tenemos la representación

$$D(t) = \sum_{i=0}^{[t/h]} X_i,$$

y así

$$E[D(t)] = \sum_{i=0}^{[t/h]} E[X_i],$$

[1] $[t/h]$ representa el mayor entero menor o igual a t/h.

Ahora
$$E[X_i] = \sum_x x P(X_i = x),$$
pero sabemos que
$$\begin{aligned} P(X_i = x) &= P(X_i = x|\text{No hubo shock})P(\text{No hubo shock}) \\ &+ P(X_i = x|\text{Hubo un shock})P(\text{Hubo un shock}) \\ &+ P(X_i = x|\text{Hubo más de un shock})P(\text{Hubo más de un shock}) \end{aligned}$$
por lo tanto
$$\begin{aligned} E[X_i] &= \sum_x x\,[P(X_i = x|\text{No hubo shock})P(\text{No hubo shock}) \\ &+ P(X_i = x|\text{Hubo un shock})P(\text{Hubo un shock}) \\ &+ P(X_i = x|\text{Hubo más de un shock})P(\text{Hubo más de un shock})] \\ &= E[X_i|\text{No hubo shock}]P(\text{No hubo shock}) \\ &+ E[X_i|\text{Hubo un shock}]P(\text{Hubo un shock}) \\ &+ E(X_i|\text{Hubo más de un shock}]P(\text{Hubo más de un shock}) \\ &= E\left[De^{-\alpha(t-L_i)}\right](\lambda h + o(h)) + o(h) \\ &= \lambda h E\left[De^{-\alpha(t-L_i)}\right] + o(h) \end{aligned}$$
Donde L_i es el tiempo en el que llega el shock en el intervalo I_i.
Por lo tanto
$$\begin{aligned} E[D(t)] &= \sum_{i=0}^{[t/h]} \left(\lambda h E\left[De^{-\alpha(t-L_i)}\right] + o(h)\right) \\ &= \lambda E[D] E\left[\sum_{i=0}^{[t/h]} h e^{-\alpha(t-L_i)}\right] + \left[\frac{t}{h}\right] o(h) \end{aligned}$$
Pero como $L_i \in I_i$, se sigue que, al hacer $h \to 0$,
$$\sum_{i=0}^{[t/h]} h e^{-\alpha(t-L_i)} \to \int_0^t e^{-\alpha(t-y)} dy.$$

Con lo cual llegamos a que

$$E[D(t)] = \frac{\lambda E[D]}{\alpha}\left(\frac{1-e^{-\alpha t}}{\alpha}\right),$$

que es la misma que habíamos calculado anteriormente.

1.6 El período ocupado $M/G/1$.

Consideremos el sistema de cola, conocido como $M/G/1$, en el cual los clientes llegan de acuerdo a un proceso Poisson con tasa λ. A su llegada, cada cliente es atendido inmediatamente o se une a la cola. Los tiempos sucesivos de servicio son independientes e idénticamente distribuidos de acuerdo con G, y son también independientes del proceso de llegada. Cuando un cliente llega y encuentra al servidor libre, decimos que comienza un período ocupado. Termina cuando no hay clientes esperando ser atendidos. Quisiéramos calcular la distribución de la longitud del período ocupado.

Supongamos que un período ocupado ha iniciado en algún momento, el cual designaremos como tiempo 0. Sea S_k el tiempo en el que llega el k-ésimo cliente adicional (así, en principio, $S_k \sim Gama(k,\lambda)$). Además, sean $Y_1, Y_2, \ldots$ la sucesión de tiempos de servicio. El período ocupado terminará al tiempo t y consistirá de n servicios si y sólo si:

i) Hay $(n-1)$ llegadas en el intervalo $(0,t)$.

ii)$Y_1 + \cdots + Y_n = t$.

iii)$S_k \leq Y_1 + \cdots + Y_k, \qquad k = 1, \ldots, n-1.$

Justificación: Supongamos que $S_k > Y_1 + \cdots + Y_k$, entonces el k-ésimo cliente adicional encontraría vacío el sistema, pues los $k-1$ anteriores habrían finalizado su servicio y, por lo tanto, el periodo ocupado termina cuando tal cliente finaliza su servicio. Lo cual es una contradicción. De la consideración anterior tenemos que

$$\begin{aligned}
& P(\text{Período ocupado tiene longitud } t \text{ y consta de } n \text{ servicios}) \\
= \;& P(Y_1 + \cdots + Y_n = t; (n-1) \text{ llegadas en } (0,t); S_k \leq Y_1 + \cdots + Y_k, \\
& k = 1, \ldots, n-1)
\end{aligned}$$

$$= \quad P(S_k \le Y_1 + \cdots + Y_k, k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t;$$
$$n-1 \text{ llegadas en } (0,t)) \times P(Y_1 + \cdots + Y_n = t; n-1 \text{ llegadas en } (0,t))$$

Ahora, como el proceso de llegadas es independiente del tiempo de servicio,

$$P(n-1 \text{ llegadas en } (0,t), Y_1 + \cdots + Y_n = t) = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}dG_n(t).$$

donde G_n es la n-ésima convolución [3] de G consigo misma. En suma, dado que hay $n-1$ llegadas en el intervalo $(0,t)$, los tiempos de llegada se distribuyen como $n-1$ variables aleatorias uniformes ordenadas y, por tanto,

$$P(\text{Período ocupado tiene longitud } t \text{ y consta de } n \text{ servicios}) \tag{4}$$
$$= \quad e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}dG_n(t)P(\tau_k \le Y_1 + \cdots + Y_k; k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t)$$

Donde $\tau_1, \ldots, \tau_{n-1}$ son independientes de $\{Y_1, \ldots, Y_n\}$ y representan los valores ordenados de un conjunto de $n-1$ variables uniformes$(0,t)$.

Para calcular la probabilidad restante, necesitamos antes de algunos lemas.

Lema 1.6.1
Sean $Y_1, \ldots, Y_n$ variables aleatorias independientes e idénticamente distribuidas no negativas, entonces

$$E[Y_1 + \cdots + Y_k | Y_1 + \cdots + Y_n = y] = \frac{k}{n}y, \qquad k = 1, \ldots, n.$$

Demostración:

$$E[Y_1 + \cdots + Y_k | Y_1 + \cdots + Y_n = y] = kE[Y_1 | Y_1 + \cdots + Y_n = y]$$

pero

$$\begin{aligned} nE[Y_1 | Y_1 + \cdots + Y_n = y] &= E[Y_1 + \cdots + Y_n | Y_1 + \cdots + Y_n = y] \\ &= y. \end{aligned}$$

Por lo tanto

$$E[Y_1 | Y_1 + \cdots + Y_n = y] = \frac{y}{n}.$$

[3] La definición de este concepto se puede encontrar en [7], Capítulo 4, (4.8).

Así finalmente

$$E[Y_1 + \cdots + Y_k | Y_1 + \cdots + Y_n = y] = \frac{k}{n} y.$$

□

Lema 1.6.2
Denotemos por $\tau_1, \ldots, \tau_n$ los valores ordenados de un conjunto de n variables aleatorias uniformes$(0,t)$. Sean $Y_1, Y_2, \ldots,$ variables aleatorias independientes e idénticamente distribuidas no negativas que también son independientes de $\{\tau_1, \ldots, \tau_n\}$, entonces

$$P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n | Y_1 + \cdots + Y_n = y) = \begin{cases} 1 - y/t & \text{si } 0 < y < t \\ 0 & \text{en otro caso.} \end{cases}$$

Demostración:
La prueba es por inducción sobre n.
Cuando $n = 1$ debemos calcular $P(Y_1 < \tau_1 | Y_1 = y)$ con $\tau_1 \sim uniforme(0,t)$, pero

$$P(Y_1 < \tau_1 | Y_1 = y) = P(y < \tau_1) = 1 - \frac{y}{t}, \qquad 0 < y < t.$$

Supongamos que el resultado es válido para $n-1$, es decir,

$$P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n-1 | Y_1 + \cdots + Y_{n-1} = y) = \begin{cases} 1 - y/t & \text{si } 0 < y < t \\ 0 & \text{en otro caso.} \end{cases}$$

Si $y \geq t$ el resultado es obvio. Supongamos que $y < t$. Para hacer uso de la hipótesis de inducción condicionaremos los valores de $Y_1 + \cdots + Y_{n-1}$ y τ_n y haremos uso del hecho de que, condicionado a que $\tau_n = u$, $\tau_1, \ldots, \tau_{n-1}$ se distribuyen como las estadísticas de orden de un conjunto de $n-1$ variables uniformes$(0,u)$. Haciendo esto para $s < y$

$$P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n | Y_1 + \cdots + Y_{n-1} = s, \tau_n = u, Y_1 + \cdots + Y_n = y)$$
$$= \begin{cases} P(Y_1 + \cdots + Y_k < \tau_k^*, k = 1, \ldots, n-1 | Y_1 + \cdots + Y_{n-1} = s) & \text{si } y < u \\ 0 & \text{en otro caso.} \end{cases}$$

Donde $\tau_1^*, \ldots, \tau_{n-1}^*$ son los valores ordenados de un conjunto de $n-1$ variables aleatorias uniformes$(0,u)$. Así por hipótesis de inducción, el lado derecho de la igualdad anterior es igual a

$$\begin{cases} 1 - s/u & \text{si } y < u \\ \\ 0 & \text{en otro caso.} \end{cases}$$

Así para $y < u$

$$\begin{aligned} & P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n | Y_1 + \cdots + Y_{n-1}, \tau_n = u, Y_1 + \cdots + Y_n = y) \\ = & \; 1 - \frac{Y_1 + \cdots + Y_{n-1}}{\tau_n}. \end{aligned} \tag{5}$$

Notemos que

$$\begin{aligned} & P(Y_1 + \cdots + Y_k < \tau_k, | Y_1 + \cdots + Y_{n-1}, \tau_n = u, Y_1 + \cdots + Y_n = y) \\ = & \; E\left[1_{\{Y_1 + \cdots + Y_k < \tau_k\}} | Y_1 + \cdots + Y_{n-1}, \tau_n = u, Y_1 + \cdots + Y_n = y\right]. \end{aligned}$$

Al tomar esperanzas condicionales en ambos lados de la igualdad (5) y considerando que

$$\sigma(\tau_n = u, Y_1 + \cdots + Y_n = y) \quad \subset \quad \sigma(\tau_n = u, Y_1 + \cdots + Y_{n-1} Y_1 + \cdots + Y_n = y),$$

obtenemos

$$\begin{aligned} & E\left[1_{\{Y_1 + \cdots + Y_k < \tau_k\}} | \tau_n = u, Y_1 + \cdots + Y_n = y\right]. \\ = & \; E\left[1 - \frac{Y_1 + \cdots + Y_{n-1}}{\tau_n} \middle| Y_1 + \cdots + Y_{n-1}, \tau_n = u, Y_1 + \cdots + Y_n = y\right]. \end{aligned}$$

Por lo tanto

$$\begin{aligned} & P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n | \tau_n = u, Y_1 + \cdots + Y_n = y) \\ = & \; E\left[1 - \frac{Y_1 + \cdots + Y_{n-1}}{\tau_n} \middle| Y_1 + \cdots + Y_{n-1}, \tau_n = u, Y_1 + \cdots + Y_n = y\right] \\ = & \; 1 - \frac{1}{u} E[Y_1 + \cdots + Y_{n-1} | Y_1 + \cdots + Y_n = y] \\ = & \; 1 - \frac{(n-1)y}{nu}. \qquad \text{Por el lema 1.6.1.} \end{aligned}$$

Tomando esperanzas un vez más tenemos que

$$
\begin{aligned}
& P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n | Y_1 + \cdots + Y_n = y) \\
= \; & E\left[1 - \frac{n-1}{n} \frac{y}{\tau_n} \middle| \tau_n > y \right] P(\tau_n > y) \\
= \; & P(\tau_n > y) - \frac{y(n-1)}{n} E\left[\frac{1}{\tau_n} \middle| \tau_n > y \right] P(\tau_n > y).
\end{aligned}
$$

Calculemos la distribución de τ_n.

$$
\begin{aligned}
P(\tau_n < x) &= P\left(\max_{1 \leq i \leq n} U_i < x \right) = P(U_1 < x, \ldots, U_n < x) \\
&= \left(\frac{x}{t} \right)^n, \qquad 0 < x < t.
\end{aligned}
$$

Donde $U_i, i = 1, \ldots, n$ son variables aleatorias uniformes$(0, t)$. Por lo tanto, su densidad está dada por

$$
f_{\tau_n}(x) = \left(\frac{n}{t} \right) \left(\frac{x}{t} \right)^{n-1} \qquad 0 < x < t,
$$

y así

$$
\begin{aligned}
E\left[\frac{1}{\tau_n} \middle| \tau_n > y \right] P(\tau_n > y) &= \int_y^t \left(\frac{1}{x} \right) \left(\frac{n}{t} \right) \left(\frac{x}{t} \right)^{n-1} dx \\
&= \frac{n}{n-1} \left(\frac{t^{n-1} - y^{n-1}}{t^n} \right).
\end{aligned}
$$

Finalmente, cuando $y < t$

$$
\begin{aligned}
P(Y_1 + \cdots + Y_k < \tau_k, \quad & k = 1, \ldots, n | Y_1 + \cdots + Y_n - y) \\
&= 1 - \left(\frac{y}{t} \right)^n - \frac{y(t^{n-1} - y^{n-1})}{t^n} \\
&= \frac{t^n - y t^{n-1}}{t^n} \\
&= 1 - \frac{y}{t}.
\end{aligned}
$$

Con lo cual concluimos la demostación.

□

Solamente necesitamos un lema más antes de regresar al problema del período ocupado.

Lema 1.6.3

Denotemos por $\tau_1, \ldots, \tau_n$ los valores ordenados de un conjunto de $n-1$ variables aleatorias uniformes $(0,t)$, y sean $Y_1, Y_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas que también son independientes de $\{\tau_1, \ldots, \tau_n\}$, entonces

$$P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) = \frac{1}{n}.$$

Demostración:

Para calcular la probabilidad anterior, haremos uso del lema 1.6.2 condicionando sobre $Y_1 + \cdots + Y_{n-1}$, esto es,

$$\begin{aligned}
& P(Y_1 + \cdots + Y_k < \tau_k,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_{n-1} = y, Y_1 + \cdots + Y_n = t) \\
= & \frac{P(Y_1 + \cdots + Y_k < \tau_k,\ k = 1, \ldots, n-1, Y_1 + \cdots + Y_{n-1} = y,\ Y_1 + \cdots + Y_n = t)}{P(Y_1 + \cdots + Y_{n-1} = y,\ Y_1 + \cdots + Y_n = t)} \\
= & \frac{P(Y_1 + \cdots + Y_k < \tau_k,\ k = 1, \ldots, n-1,\ Y_1 + \cdots + Y_{n-1} = y)P(Y_n = t - y)}{P(Y_1 + \cdots + Y_{n-1} = y)P(Y_n = t - y)} \\
= & P(Y_1 + \cdots + Y_k < \tau_k,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_{n-1} = y) \\
= & \begin{cases} 1 - y/t \text{ si } 0 < y < t \\ \\ 0 \text{ en otro caso} \end{cases}
\end{aligned}$$

Ahora, como $Y_1 + \cdots + Y_{n-1} \leq Y_1 + \cdots + Y_n$ tenemos que

$$\begin{aligned}
& P(Y_1 + \cdots + Y_k < \tau_k,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= & E\left[1 - \left(\frac{Y_1 + \cdots + Y_{n-1}}{t}\right) \middle| Y_1 + \cdots + Y_n = t\right] \\
= & 1 - \left(\frac{n-1}{n}\right) \qquad \text{Por el lema 1.6.1} \\
= & \frac{1}{n}.
\end{aligned}$$

Por lo tanto

$$P(Y_1 + \cdots + Y_k < \tau_k, k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) = \frac{1}{n}.$$

□

Regresando a la distribución conjunta de la longitud de un periodo ocupado y el número de clientes atendidos que dejamos pendiente en la página 28, debemos calcular

$$P(\tau_k \leq Y_1 + \cdots + Y_k,\ k = 1, \ldots, n-1 \mid Y_n = t).$$

Ahora, dado que, si U es una variable aleatoria Uniforme $(0,t)$, $t - U$ también lo es, se sigue que $(\tau_1, \ldots, \tau_{n-1})$ tiene la misma distribución conjunta que $(t - \tau_{n-1}, \ldots, \tau_1)$. De aquí reemplazando τ_k por $t - \tau_{n-k}$, $k = 1, \ldots, n-1$ obtenemos

$$\begin{aligned}
& P(\tau_k \leq Y_1 + \cdots + Y_k,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & P(t - \tau_{n-k} \leq Y_1 + \cdots + Y_k,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & P(t - \tau_{n-k} \leq t - (Y_{k+1}, \ldots, Y_n),\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & P(\tau_{n-k} \geq Y_{k+1}, \ldots, Y_n,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & P(\tau_{n-k} \geq Y_{n-k}, \ldots, Y_1,\ k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t)
\end{aligned}$$

Donde la última igualdad se obtiene del hecho de que $Y_1 + \cdots + Y_n$ tiene la misma distribución conjunta que $Y_n, \ldots, Y_1$, y así , cualquier probabilidad que relacione las Y_i's sigue siendo válida si Y_1 es reemplazada por Y_n, Y_2 por Y_{n-1}, Y_k por $Y_{n-k+1}, \ldots, Y_1$ por Y_n. De aquí tenemos que

$$\begin{aligned}
& P(\tau_k \leq Y_1 + \cdots + Y_k, k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & P(\tau_k \geq Y_1 + \cdots + Y_k, k = 1, \ldots, n-1 \mid Y_1 + \cdots + Y_n = t) \\
= \ & \frac{1}{n} \qquad \text{Por el lema 1.6.3.}
\end{aligned}$$

Así si denotamos $B(t,n)$ = P(periodo ocupado tenga longitud menor o igual a t y n clientes atendidos en un periodo ocupado) entonces

$$\frac{d}{dt}B(t,n) = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}dG_n(t)$$

o equivalentemente

$$B(t,n) = \int_0^t e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}dG_n(t).$$

Finalmente, si $B(t)$ representa la distribución de la longitud de un periodo ocupado, entonces

$$B(t) = \sum_{n=1}^{\infty} B(t,n),$$

está dada por

$$B(t) = \sum_{n=1}^{\infty} \int_0^t e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} dG_n(t)$$

1.7 Proceso Poisson no homogéneo.

En esta sección generalizaremos el proceso Poisson al permitir que la tasa de llegada sea una función del tiempo.

Definición 1.7.1

El proceso de conteo $\{N(t), t \geq 0\}$ es llamado Proceso Poisson no homogéneo o no estacionario con función de intensidad $\lambda(t)$ si:

i)$N(0) = 0$.

ii)$\{N(t), t \geq 0\}$ *tiene incrementos independientes.*

iii)$P(N(t+h) - N(t) \geq 2) = o(h)$.

iv)$P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$.

Definamos

$$m(t) = \int_0^t \lambda(s)ds.$$

Mostraremos que

$$P(N(t+s) - N(t) = n) = \exp\{-[m(t+s) - m(t)]\} \frac{[m(t+s) - m(t)]^n}{n!}; \quad n \geq 0.$$

Es decir, $N(t+s) - N(t)$ tiene una distribución Poisson con media $m(t+s) - m(t)$.

Demostración:

Sea $t \geq 0$ y definamos $P_n(u) = P(N(t+u) - N(t) = n)$.

$$P_0(s+h) \;=\; P(N(t+s+h) - N(t) = 0)$$

$$\begin{aligned}
&= P(N(t+s)-N(t)=0, N(t+s+h)-N(t+s)=0) \\
&= P(N(t+s)-N(t)=0)P(N(t+s+h)-N(t+s)=0) \\
&= P_0(s)P(N(t+s+h)-N(t+s)=0) \\
&= P_0(s)[1-P(N(t+s+h)-N(t+s)=1) \\
&- P(N(t+s+h)-N(t+s)\geq 2)] \\
&= P_0(s)[1-\lambda(t+s)h+o(h)] \\
&= P_0(s)-\lambda(t+s)hP_0(s)+o(h)
\end{aligned}$$

Así

$$P_0(s+h)-P_0(s) = -\lambda(t+s)hP_0(s)+o(h)$$

y al dividir entre h y hacer $h \to 0$:

$$P_0'(s) = -\lambda(t+s)P_0(s).$$

cuya solución con condición inicial $P_0(0)=1$ es

$$P_0(s) = \exp\left\{-\int_0^s \lambda(t+u)du\right\}.$$

Notemos que

$$\begin{aligned}
\int_0^s \lambda(t+u)du &= \int_t^{t+s} \lambda(w)dw = \int_0^{t+s} \lambda(w)dw - \int_0^t \lambda(w)dw \\
&= m(t+s)-m(t).
\end{aligned}$$

Por lo tanto

$$P_0(s) = \exp\{-[m(t+s)-m(t)]\}.$$

Ahora obtendremos una igualdad semejante para $P_n(s)$. Dado que

$$\begin{aligned}
P_n(s+h) &= P(N(t+s+h)-N(t)=n) \\
&= P(N(t+s)-N(t)=n,\ N(t+s+h)-N(t+s)=0) \\
&+ P(N(t+s)-N(t)=n-1,\ N(t+s+h)-N(t+s)=1) \\
&+ P(N(t+s+h)-N(t)=n,\ N(t+s+h)-N(t+s)\geq 2)
\end{aligned}$$

tenemos que

$$\begin{aligned} P_n(s+h) &= P_n(s)P(N(t+s+h)-N(t)=0) \\ &+ P_{n-1}(s)P(N(t+s+h)-N(t+s)=1)+o(h). \end{aligned}$$

Al simplificar llegamos a que

$$\frac{P_n(s+h)-P_n(s)}{h} = -\lambda(t+s)P_n(s)+\lambda(t+s)P_{n-1}(s)+\frac{o(h)}{h}$$

y haciendo $h \to 0$ y multiplicando ambos lados de la igualdad por $\exp\{\int_0^s \lambda(t+u)du\}$ tenemos que

$$\left(P_n(s)\exp\left\{\int_0^s \lambda(t+u)du\right\}\right)' = \lambda(t+s)P_{n-1}(s)\exp\left\{\int_0^s \lambda(t+u)du\right\}$$

o equivalentemente

$$(P_n(s)e^{m(t+s)-m(t)})' = \lambda(t+s)P_{n-1}(s)\exp\{m(t+s)-m(t)\}.$$

Para comprobar la igualdad, hagamos inducción sobre n

Evaluando en $n=1$ con condiciones iniciales $P_1(0)=0$ tenemos

$$P_1(s) = [m(t+s)-m(t)]e^{-[m(t+s)-m(t)]},$$

es decir, se cumple la igualdad para $n=1$.

Ahora supongamos que la igualdad es válida para $n-1$, es decir

$$P_{n-1}(s) = e^{-[m(t+s)-m(t)]}\frac{[m(t+s)-m(t)]^{n-1}}{(n-1)!}$$

Así

$$\begin{aligned} (P_n(s)e^{m(t+s)-m(t)})' &= \lambda(t+s)\frac{[m(t+s)-m(t)]^{n-1}}{(n-1)!}e^{-[m(t+s)-m(t)]} \\ &\times\ e^{m(t+s)-m(t)} \\ &= \lambda(t+s)\frac{[m(t+s)-m(t)]^{n-1}}{(n-1)!} \end{aligned}$$

o equivalentemente

$$P_n(s)e^{m(t+s)-m(t)} = \int_0^s \lambda(t+u)\frac{[m(t+u)-m(t)]^{n-1}}{(n-1)!}du.$$

Finalmente, al resolver la integral haciendo el cambio de variable

$$\begin{aligned} v &= m(t+u) - m(t) \qquad \text{y} \\ dv &= \lambda(t+u)du \end{aligned}$$

tenemos que

$$P(N(t+s) - N(t) = n) = e^{-[m(t+s)-m(t)]}\frac{[m(t+s) - m(t)]^n}{n!}.$$

Lo cual concluye la demostración.

□

Así $N(t+s) - N(t)$ efectivamente se distribuye Poisson con media $m(t+s) - m(t)$. La importancia del Proceso Poisson no homogéneo reside en el hecho de que, al no requerir incrementos estacionarios, damos la posibilidad de que los eventos ocurran con mayor probabilidad en un instante dado que en otro.

Cuando la función de intensidad $\lambda(t)$ está acotada, podemos pensar al proceso Poisson no homogéneo como una muestra aleatoria de un proceso Poisson homogéneo.

Especificamente, sea λ tal que

$$\lambda(t) \leq \lambda \qquad \forall t \geq 0$$

y consideremos el proceso Poisson con tasa λ. Si suponemos que un evento del proceso Poisson que ocurre al tiempo t es incluido en la muestra (contado) con probabilidad $\lambda(t)/\lambda$, entonces el proceso de inclusión en la muestra es un proceso Poisson no homogéneo con función de intensidad $\lambda(t)$.

Efectivamente, los incisos $i), ii)$ y $iii)$ de la definición 1.7.1 se cumplen, pues son también ciertos para el proceso Poisson.

El axioma $iv)$ se sigue de que

$$\begin{aligned} & P(\text{Hay un evento contado en el intervalo } (t, t+h)) \\ = \; & P(\text{Hay un evento en el intervalo } (t, t+h) \text{ y contarlo}) \\ + \; & P(\text{Hay dos o más un eventos en } (t, t+h) \text{ y contar sólo uno}) \\ = \; & P(\text{Hay un evento en el intervalo } (t, t+h))\frac{\lambda(t)}{\lambda} + o(h) \\ = \; & (\lambda h + o(h))\frac{\lambda(t)}{\lambda} + o(h) \\ = \; & \lambda(t)h + o(h). \end{aligned}$$

La interpretación de un proceso Poisson no homogéneo como una muestra de uno homogéneo también nos da otra manera de entender la proposición 1.5.1 de la página 21.

Ejemplo 1.7.1

Sea $X_1, X_2, \ldots,$ una sucesión de variables aleatorias independientes e identicamente distribuidas no negativas y continuas cuya función de intensidad está dada por $\lambda(t) = f(t)/(1 - F(t))$, donde f y F son respectivamente la función de densidad y distribución. Decimos que un registro ocurre al tiempo n si $X_n > \max\{X_1, \ldots, X_{n-1}\}$ donde $X_0 = 0$. Si un registro ocurre al tiempo n, entonces X_n es llamado un valor de registro.

Sea $N(t)$ el número de valores extremos (respecto a su pasado) menores o iguales que t. Esto es, $N(t)$ es un proceso de conteo de eventos, donde un evento ocurre al tiempo x si x es un valor extremo.

Afirmamos que $\{N(t), t \geq 0\}$ será un proceso Poisson no homogéneo con intensidad $\lambda(t)$. Para verificar esta afirmación notemos que habrá un valor de registro entre t y $t+h$ si y sólo si la primera X_i cuyo valor es más grande que t cae entre t y $t+h$. Supongamos que X_n es el primer valor de registro que cumple la condición anterior, entonces

$$\begin{aligned}
P(N(t+h) - N(t) = 1) &= P(X_n \in (t, t+h) \mid X_n > t) \\
&= \frac{P(X_n \in (t, t+h), X_n > t)}{P(X_n > t)} \\
&= \frac{f(t)h + o(h)}{1 - F(t)} \\
&= \frac{f(t)}{1 - F(t)} h + o(h) \\
&= \lambda(t)h + o(h)
\end{aligned}$$

Los incisos i) y ii) son inmediatos y el iii) se sigue del iv).

1.8 Proceso Poisson compuesto.

Definición 1.8.1
El proceso estocástico $\{X(t), t \geq 0\}$ es llamado proceso Poisson compuesto si puede ser representado como

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \qquad t \geq 0.$$

Donde $\{N(t), t \geq 0\}$ es un proceso Poisson y $\{Y_i, i = 1, 2, \ldots\}$ es una familia de variables aleatorias independientes e idénticamente distribuidas. El proceso Poisson $\{N(t), t \geq 0\}$ y la sucesión $\{Y_i, i = 1, 2, \ldots\}$ son independientes.

Como un ejemplo de un proceso Poisson compuesto supongamos que la llegada de los clientes a una tienda se realiza de acuerdo a un proceso Poisson con tasa λ. Supongamos, también que la cantidad de dinero gastado por cada cliente forma un conjunto de variables aleatorias independientes e identicamente distribuidas, entonces, si $X(t)$ denota la cantidad total de dinero gastada en la tienda al tiempo t, vemos que $\{X(t), t \geq 0\}$ es un proceso Poisson compuesto.
Calculemos la función generadora de momentos de $X(t)$

$$\begin{aligned}
\phi_{X(t)}(u) &= E[\exp\{uX(t)\}] \\
&= \sum_{n=0}^{\infty} E[\exp\{uX(t)\} \mid N(t) = n] P[N(t) = n] \\
&= \sum_{n=0}^{\infty} E[\exp\{u(Y_1 + \cdots Y_n)\} \mid N(t) = n] P[N(t) = n] \\
&= \sum_{n=0}^{\infty} E[\exp\{u(Y_1 + \cdots Y_n)\} \mid N(t) = n] \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= \sum_{n=0}^{\infty} E[\exp\{u(Y_1)\}]^n \frac{e^{-\lambda t}(\lambda t)^n}{n!}
\end{aligned}$$

Sea $\phi_Y(u) = E[\exp\{u(Y)\}]$, así

$$\begin{aligned}
\phi_{X(t)}(u) &= \sum_{n=0}^{\infty} (\phi_Y(u))^n \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= \exp\{\lambda t(\phi_Y(u) - 1)\}.
\end{aligned}$$

Diferenciando, es fácil comprobar que

$$\begin{aligned} E[X(t)] &= \lambda t E[Y] \\ Var[X(t)] &= \lambda t E[Y^2]. \end{aligned}$$

1.9 Proceso Poisson condicional.

Definición 1.9.1

Sean Λ una variable aleatoria positiva con distribución G y $\{N(t), t \geq 0\}$ un proceso de contar tal que, dado que $\Lambda = \lambda$, $\{N(t), t \geq 0\}$ es un proceso Poisson con tasa λ. El proceso $\{N(t), t \geq 0\}$ es llamado proceso Poisson condicional.

Así en principio

$$\begin{aligned} P(N(t+s) - N(s)) &= \int_0^\infty P(N(t+s) - N(s), \Lambda = \lambda) d\lambda \\ &= \int_0^\infty P(N(t+s) - N(s) \mid \Lambda = \lambda) g(\lambda) d\lambda \\ &= \int_0^\infty P(N(t+s) - N(s) \mid \Lambda = \lambda)) dG(\lambda) \\ &= \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^n}{n!} dG(\lambda). \end{aligned}$$

Es importante hacer notar que $\{N(t), t \geq 0\}$ no es propiamente un proceso Poisson, pues aunque tiene incrementos estacionarios, éstos no son independientes.

Para comprobarlo, sean $t_{i-1} < t_i < t_{j-1} < t_j$

$$\begin{aligned} & P(N(t_i) - N(t_{i-1}) = n, N(t_j) - N(t_{j-1}) = m) \\ = & \int_0^\infty P(N(t_i) - N(t_{i-1}) = n, N(t_j) - N(t_{j-1}) = m \mid \Lambda = \lambda) dG(\lambda). \end{aligned}$$

Pero sabemos que condicionado a que $\Lambda = \lambda$, $\{N(t), t \geq 0\}$ sí es un proceso Poisson, por lo tanto tiene incrementos estacionarios e independientes, de aquí que

$$\begin{aligned} & P(N(t_i) - N(t_{i-1}) = n, N(t_j) - N(t_{j-1}) = m) \\ = & \int_0^\infty \left\{ \frac{e^{-\lambda(t_i - t_{i-1})}[\lambda(t_i - t_{i-1}]^n}{n!} \right\} \left\{ \frac{e^{-\lambda(t_j - t_{j-1})}[\lambda(t_j - t_{j-1}]^m}{m!} \right\} dG(\lambda) \\ \neq & P(N(t_i) - N(t_{i-1}) = n) P(N(t_j) - N(t_{j-1}) = m). \end{aligned}$$

Así efectivamente $\{N(t), t \geq 0\}$ no tiene incrementos independientes.

Ahora calculemos la distribución condicional de Λ dado que $N(t) = n$. Para $d(\lambda)$ pequeño:

$$
\begin{aligned}
P(\Lambda \in (\lambda, \lambda + d\lambda) \mid N(t) = n) &= \frac{P(N(t) = n \mid \Lambda \in (\lambda, \lambda + d\lambda))P(\Lambda \in (\lambda, \lambda + d\lambda)}{P(N(t) = n)} \\
&= \frac{\frac{e^{-\lambda t}(\lambda t)^n}{n!}dG(\lambda)}{\int_0^\infty \frac{e^{-\lambda t}(\lambda t)^n}{n!}dG(\lambda)}.
\end{aligned}
$$

Y así la distribución condicional de Λ, dado que $N(t) = n$, está dada por

$$
P(\Lambda \leq x \mid N(t) = n) = \frac{\int_0^x \frac{e^{-\lambda t}(\lambda t)^n}{n!}dG(\lambda)}{\int_0^\infty \frac{e^{-\lambda t}(\lambda t)^n}{n!}dG(\lambda)}.
$$

Capítulo 2

El método de Chen-Stein.

2.1 Una idea global.

En 1972, Charles Stein Publicó un artículo llamado *"A bound on the error in the normal approximation to the distribution of a sum of dependent random variables"* [14]. El mérito de este trabajo fue mostrar la convergencia en distribución a la Normal estandar y dar cotas para el error hecho en la aproximación sin hacer uso de metodos de Fourier.

Stein utilizó la ecuación diferencial

$$f'(x) - xf(x) = h(x) - Nh$$

donde:
f es una función diferencial.
h es una función que se utiliza para la prueba de la convergencia en distribución.
$Nh = E[h(Z)]$ con $Z \sim Normal(0,1)$.

La conexión entre esta ecuación y la distribución normal es la siguiente caracterización: Para una variable aleatoria arbitraria W y

$$(Lf)(x) = f'(x) - xf(x),$$

$E[Lf(W)] = 0$ para toda función diferenciable f tal que $E[\mid Zf(Z) \mid] < \infty$ si y sólo si $W \sim Normal(0,1)$.

Parece razonable que si $E[Lf(W)]$ es pequeña para muchas funciones f, entonces la distribución de W es cercana a la de Z.

En 1975, L. Chen, quien fuera alumno de Stein, aplicó la idea de este último al caso Poisson obteniendo una ecuación en diferencias análoga a la ecuación de Stein con $Z \sim Poisson(\lambda)$, si definimos

$$(Lf)(x) = \lambda f(x+1) - xf(x)$$

entonces $E[Lf(W)] = 0$ para toda función f tal que $E[\| Zf(Z) \|] < \infty$ si y sólo si $W \sim Poisson(\lambda)$.

Análogamente, parece razonable creer que si $E[Lf(W)]$ es pequeña para muchas funciones f, la distribución de W es parecida a la de Z.

A continuación, definiremos y analizaremos las propiedades de algunos operadores que serán de suma utilidad en la demostracion del teorema principal de este trabajo y que enunciaremos un poco más adelante.

2.2 Definición de un operador inverso.

Recordemos que $Z \sim Poisson(\lambda)$. Definamos los operadores lineales S y T, que dependen del parámetro λ, por

$$\begin{aligned} (Tf)(w) &\equiv wf(w) - \lambda f(w+1), && \text{para } w \geq 0 \\ (Sh)(w+1) &\equiv -\lambda^{-1}P(Z=w)^{-1}E[h(Z); Z \leq w], && \text{para } w \geq 0. \end{aligned}$$

Para estar completamente definido, elegimos $(Sh)(0) = 0$, la cual es una elección arbitraria pues el valor de $(Sh)(0)$ nunca se utiliza.

Notemos que S es inversa a T para toda h pues $T(Sh) = h$. En efecto,

Sea

$$\mathcal{E}_w = E[h(Z); Z \leq w] = \sum_{n=0}^{w} h(n)P(Z=n).$$

Sea $w \geq 0$ arbitrario.

$$(T(Sh))(w+1) \;=\; (w+1)Sh(w+1) - \lambda Sh(w+2)$$

$$\begin{aligned}
&= (w+1)\left\{-\left(\frac{1}{\lambda}\right)e^{\lambda}\frac{w!}{\lambda^{w}}\mathcal{E}_w\right\}-\lambda\left\{-\left(\frac{1}{\lambda}\right)e^{\lambda}\frac{(w+1)!}{\lambda^{w+1}}\mathcal{E}_{w+1}\right\}\\
&= \left(\frac{e^{\lambda}(w+1)!}{\lambda^{w+1}}\right)(\mathcal{E}_{w+1}-\mathcal{E}_w)\\
&= P(Z=w+1)^{-1}h(w+1)P(Z=w+1)\\
&= h(w+1).
\end{aligned}$$

Por lo tanto $(T(Sh))(w+1)=h(w+1)$, lo que muestra que $T(Sh)=h$.

Además, para toda función f acotada se tiene que $E[(Tf)(Z)]=0$, lo cual es facilmente verificable:

Como f es acotada existe $M\in\mathbb{R}$ tal que $f(x)\le M$ para toda x que esté en el dominio de f, por lo tanto

$$E[Zf(Z)]\le E[ZM]=M\lambda<\infty \quad \text{y} \quad E[(Tf)(Z)]\le\lambda M<\infty.$$

De aquí que

$$E[(Tf)(Z)]=E[Zf(Z)]-E[\lambda f(Z+1)].$$

Ahora:

$$\begin{aligned}
& E[(Tf)(Z)]=0 \iff E[Zf(Z)]-E[\lambda f(Z+1)]=0\\
\iff\; & E[Zf(Z)] = E[\lambda f(Z+1)]\\
\iff\; & \sum_{k=0}^{\infty}kf(k)P(Z=k) = \sum_{k=0}^{\infty}\lambda f(k+1)P(Z=k)\\
\iff\; & \sum_{k=1}^{\infty}\frac{f(k)e^{-\lambda}\lambda^{k}}{(k-1)!} = \sum_{k=0}^{\infty}\frac{f(k+1)e^{-\lambda}\lambda^{k+1}}{k!}\\
\iff\; & \sum_{k=1}^{\infty}\frac{f(k)\lambda^{k}}{(k-1)!} = \sum_{k=0}^{\infty}\frac{f(k+1)\lambda^{k+1}}{k!}
\end{aligned}$$

La última igualdad se da al hacer un cambio de variable en la última serie del lado izquierdo para obtener la del lado derecho. Por lo tanto

$$E[(Tf)(Z)]=0 \quad \forall f \text{ acotada}.$$

En la siguiente sección definiremos la notación que utilizaremos a lo largo de este trabajo y enunciaremos el primer teorema fundamental de este escrito.

2.3 Notación y teorema principal.

Sea I un conjunto arbitrario de índices, y para $\alpha \in I$, sea X_α una variable aleatoria Bernoulli con $p_\alpha \equiv P(X_\alpha = 1) = 1 - P(X_\alpha = 0), p_\alpha > 0$.

Sean

$$W \equiv \sum_{\alpha \in I} X_\alpha \qquad \text{y} \qquad \lambda \equiv E[W] = \sum_{\alpha \in I} p_\alpha.$$

Suponemos que $\lambda \in (0, \infty)$.

Para cada $\alpha \in I$, supongamos que hemos elegido $B_\alpha \subset I$, con $\alpha \in B_\alpha$, de tal manera que si $\beta \notin B_\alpha, X_\alpha$ es independiente o "casi independiente" de X_β. Decimos que B_α es una vecindad de dependencia de X_α (aunque estrictamente hablando B_α es una vecindad de α).

Definamos

$$\begin{aligned} b_1 &\equiv \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} p_\alpha p_\beta. \\ b_2 &\equiv \sum_{\alpha \in I} \sum_{\alpha \neq \beta \in B_\alpha} p_{\alpha\beta} \qquad \text{donde} \quad p_{\alpha\beta} \equiv E[X_\alpha X_\beta]. \\ b_3' &\equiv \sum_{\alpha \in I} s_\alpha' \qquad \text{y } b_3 \equiv \sum_{\alpha \in I} s_\alpha. \end{aligned}$$

donde

$$\begin{aligned} s_\alpha' &\equiv E\left[\left|E\left[(X_\alpha - p_\alpha) | \sum_{\beta \in I - B_\alpha} X_\beta\right]\right|\right] \\ &\leq s_\alpha \equiv E\left[\left|E\left[(X_\alpha - p_\alpha)\right| \sigma(X_\beta : \beta \in I - \{\alpha\})\right]\right|. \end{aligned} \tag{6}$$

La desigualdad (6) se puede verificar utilizando el hecho de que

$$\sigma\left(\sum_{\beta \in I - B_\alpha} X_\beta\right) \subset \sigma(X_\beta : \beta \in I - \{\alpha\}),$$

además de la desigualdad de Jensen y que si $\mathcal{F}_1$y$\mathcal{F}_2$ son dos σ-álgebras tales que $\mathcal{F}_1 \subset \mathcal{F}_2$ entonces:

(i)$E[E(X|\mathcal{F}_1)|\mathcal{F}_2]] = E[X|\mathcal{F}_1]$.

(i)$E[E(X|\mathcal{F}_2)|\mathcal{F}_1]] = E[X|\mathcal{F}_1]$.

En efecto, sean

$$\mathcal{F}_1 = \sigma\left(\sum_{\beta \in I - B_\alpha} X_\beta\right) \quad \text{y}$$
$$\mathcal{F}_2 = \sigma(X_\beta : \beta \in I - \{\alpha\}).$$

De aquí que

$$\begin{aligned} |E[(X_\alpha - p_\alpha) \mid \mathcal{F}_1]| &= |E[E[(X_\alpha - p_\alpha) \mid \mathcal{F}_1] \mid \mathcal{F}_2]| \\ &= |E[E[(X_\alpha - p_\alpha) \mid \mathcal{F}_2] \mid \mathcal{F}_1]| \\ &\leq E[|E[(X_\alpha - p_\alpha) \mid \mathcal{F}_1]| \mid \mathcal{F}_2] \\ &= E[|E[(X_\alpha - p_\alpha) \mid \mathcal{F}_1]|] \end{aligned}$$

A grandes rasgos, el resultado que enunciaremos en un momento, indica que cuando b_1, b_2 y b_3 son pequeños, entonces

> *El número total de eventos W tiene aproximadamente una distribución Poisson (Teorema 2.3.1).*

Se puede decir que b_1 mide el tamaño de la vecindad, b_2 mide el número esperado de vecindades de una ocurrencia dada y b_3' o b_3 mide la dependencia entre un evento y el número de ocurrencias fuera de su vecindad.

Denotemos por Z a una variable aleatoria Poisson con media λ.

Sean $f, h : \mathbb{Z}^+ \to \mathbb{R}$, donde $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ y $\|h\| = \sup_{k \geq 0} |h(k)|$.

Definamos la distancia de la variación total entre las distribuciones de W y de Z como

$$\begin{aligned} \|\mathcal{L}(W) - \mathcal{L}(Z)\| &= \sup_{\|h\|=1} |E[h(W)] - E[h(Z)]| \\ &= 2 \sup_{A \in \mathbb{Z}^+} |P(W \in A) - P(Z \in A)|. \end{aligned}$$

Observemos que la convergencia en distribución es equivalente a la convergencia bajo la métrica de Prohorov, la cual coincide con la mitad de la distancia de la variación total en el conjunto de medidas con dominio en $\mathbb{Z}^+$. La demostración de este hecho se puede encontrar en [16] .

A continuación enunciaremos el teorema principal de este trabajo.

TEOREMA 2.3.1
Sea W el número de ocurrencias de eventos dependientes, y sea Z una variable aleatoria Poisson con $E[Z] = E[W] = \lambda$. Entonces

$$
\begin{aligned}
\|\mathcal{L}(W) - \mathcal{L}(Z)\| &\leq 2\left[(b_1 + b_2)\frac{1-e^{-\lambda}}{\lambda} + b_3' \min\{1, 1.25\lambda^{-1/2}\}\right] \\
&\leq 2(b_1 + b_2 + b_3).
\end{aligned}
$$

además

$$
\begin{aligned}
|P(W=0) - e^{-\lambda}| &\leq (b_1 + b_2 + b_3')\left(1 - e^{-\lambda}\right)/\lambda. \\
&< \min 1, \lambda^{-1}(b_1 + b_2 + b_3).
\end{aligned}
$$

2.4 Algunas observaciones.

Antes de entrar en la demostración del teorema daremos algunas observaciones previas:

Sean

$$V_\alpha \equiv \sum_{\beta \in I - B_\alpha} X_\beta \qquad \text{y} \qquad W_\alpha \equiv W - X_\alpha.$$

Así

$$V_\alpha \leq W_\alpha \leq W.$$

Sea f una función arbitraria.

Observación 2.4.1

$$X_\alpha f(W) = X_\alpha f(W_\alpha + 1).$$

Demostración:

Si $X_\alpha = 0$ la igualdad es trivial.
Si $X_\alpha = 1 \Longrightarrow W = W_\alpha + X_\alpha = W_\alpha + 1$.
Por lo tanto $X_\alpha f(W) = X_\alpha f(W_\alpha + 1)$.

□

Observación 2.4.2

$$f(W_\alpha + 1) - f(W + 1) = X_\alpha[f(W_\alpha + 1) - f(W_\alpha + 2)].$$

Demostración:

Si $X_\alpha = 0$ entonces $W_\alpha = W$, por lo tanto, $W_\alpha + 1 = W + 1$. De aqui que $f(W_\alpha+1) = f(W+1)$ y así $f(W_\alpha+1) - f(W+1) = 0$ y se dá la igualdad.

Si $X_\alpha = 1$ entonces $W = W_\alpha + X_\alpha = W_\alpha + 1$, por lo tanto, $W + 1 = W_\alpha + 2$ y entonces $f(W+1) = f(W_\alpha+2)$ y por tanto

$$\begin{aligned} f(W_\alpha+1) - f(W+1) &= f(W_\alpha+1) - f(W_\alpha+2) \\ &= X_\alpha[f(W_\alpha+1) - f(W_\alpha+2)]. \end{aligned}$$

□

Observación 2.4.3

$$\sum_{\alpha\in I} E\left[X_\alpha p_\alpha \left(f(W_\alpha+1) - f(W_\alpha+2)\right)\right] \leq \|\Delta f\| \sum_{\alpha\in I} p_\alpha^2.$$

Demostración:

$$\begin{aligned} f(W_\alpha+1) - f(W_\alpha+2) &= -\Delta f(W_\alpha+1), \quad \text{por lo tanto,} \\ f(W_\alpha+1) - f(W_\alpha+2) &\leq |f(W_\alpha+1) - f(W_\alpha+2)| \\ &= |\Delta f(W_\alpha+1)| \\ &\leq \sup_{k\geq 0} |\Delta f(k)| \\ &= \|\Delta f\|. \end{aligned}$$

Así $f(W_\alpha+1) - f(W_\alpha+2) \leq \|\Delta f\|$, además $E[X_\alpha] = p_\alpha \quad \forall \alpha \in I$. Por lo tanto

$$\begin{aligned} \sum_{\alpha\in I} E\left[X_\alpha p_\alpha \left(f(W_\alpha+1) - f(W_\alpha+2)\right)\right] &\leq \|\Delta f\| \sum_{\alpha\in I} p_\alpha E[X_\alpha] \\ &= \|\Delta f\| \sum_{\alpha\in I} p_\alpha^2. \end{aligned}$$

□

Observación 2.4.4

$$\sum_{\alpha\in I} E\left[(X_\alpha - p_\alpha) f(V_\alpha+1)\right] \leq b_3' \|f\|.$$

Demostración:

$$\begin{aligned}
E[(X_\alpha - p_\alpha)f(V_\alpha+1)] &= E\left[E\left[(X_\alpha - p_\alpha)f(V_\alpha+1)\middle|\sum_{\beta\in I-B_\alpha} X_\beta\right]\right] \\
&\leq E\left[\left|E\left[(X_\alpha - p_\alpha)f(V_\alpha+1)\middle|\sum_{\beta\in I-B_\alpha} X_\beta\right]\right|\right] \\
&\leq \|f\|E\left[\left|E\left[(X_\alpha - p_\alpha)\middle|\sum_{\beta\in I-B_\alpha} X_\beta\right]\right|\right] \\
&= \|f\|s'_3
\end{aligned}$$

Por lo tanto

$$\sum_{\alpha\in I} E\left[(X_\alpha - p_\alpha)f(V_\alpha+1)\right] \leq \|f\|\sum_{\alpha\in I} s'_3 = b'_3\|f\|.$$

□

Observación 2.4.5

$$\sum_{\alpha\in I} E\left[(X_\alpha - p_\alpha)[f(W_\alpha+1) - f(V_\alpha+1)]\right] \leq \|\Delta f\|\sum_{\alpha\in I}\sum_{\alpha\neq\beta\in B_\alpha}(p_{\alpha\beta} + p_\alpha p_\beta).$$

Demostración:
Sabemos que

$$\sum_{\beta\in I-B_\alpha} X_\beta = V_\alpha < W_\alpha = \sum_{\beta\in I} X_\beta - X_\alpha.$$

Por lo tanto

$$V_\alpha = W_\alpha - \sum_{\alpha\neq\beta\in B_\alpha} X_\beta,$$

es decir, podemos obtener V_α a partir de W_α restando sucesivamente $|B_\alpha| - 1$ X_β's con $\beta\in B_\alpha - \{\alpha\}$.

Ahora

$$\begin{aligned}
&E\left[(X_\alpha - p_\alpha)[f(W_\alpha+1) - f(V_\alpha+1)]\right] \\
&= E\left[(X_\alpha - p_\alpha)f(W_\alpha+1)\right] - E\left[(X_\alpha - p_\alpha)f(V_\alpha+1)\right].
\end{aligned}$$

Utilizaremos las observaciones anteriores para escribir la última expresión como una suma telescópica de $|B_\alpha| - 1$ términos. Para ello, basta decir que los términos de la suma son de la forma

$$E\left[(X_\alpha - p_\alpha)f(U + X_\beta)\right] - E\left[(X_\alpha - p_\alpha)f(U)\right]$$

donde el primer término es

$$E\left[(X_\alpha - p_\alpha)f(W_\alpha + 1)\right] - E\left[(X_\alpha - p_\alpha)f(W_\alpha + 1 - X_\beta)\right]$$

para $\beta \in B_\alpha - \{\alpha\}$, y el último es

$$E\left[(X_\alpha - p_\alpha)f(V_\alpha + 1 + X_\gamma)\right] - E\left[(X_\alpha - p_\alpha)f(V_\alpha + 1)\right]$$

para $\gamma \in B_\alpha - \{\alpha\}$ y $\gamma \neq \beta$.
Notemos que las U's tienen la forma

$$U = W_\alpha + 1 - \sum_{\beta \in B} X_\beta; \qquad \text{con } B \subset B_\alpha - \{\alpha\}.$$

Tomando en cuenta que $X_\beta = 0$ ó 1, es fácil checar que

$$\begin{aligned}
E[(X_\alpha - p_\alpha)f(U + X_\beta) - f(U)] &= E[(X_\alpha - p_\alpha)X_\beta(f(U+1) - f(U))] \\
&= E[X_\alpha X_\beta \Delta f(U)] + E[p_\alpha X_\beta(-\Delta f(U))] \\
&\leq E[X_\alpha X_\beta \|\Delta f\|] + E[p_\alpha X_\beta \|\Delta f\|] \\
&= \|\Delta f\|(p_{\alpha\beta} + p_\alpha p_\beta).
\end{aligned}$$

De aquí que

$$E[(X_\alpha - p_\alpha)(f(W_\alpha + 1) - f(V_\alpha + 1))] \leq \|\Delta f\| \sum_{\alpha \neq \beta \in B_\alpha} (p_{\alpha\beta} + p_\alpha p_\beta).$$

Por lo tanto

$$\sum_{\alpha \in I} E[(X_\alpha - p_\alpha)(f(W_\alpha + 1) - f(V_\alpha + 1))] \leq \|\Delta f\| \sum_{\alpha \in I} \sum_{\alpha \neq \beta \in B_\alpha} (p_{\alpha\beta} + p_\alpha p_\beta).$$

$\square$

Observación 2.4.6

Recordemos que

$f^+(x) = \sup\{f(x), 0\}$

$f^-(x) = \sup\{-f(x), 0\}$.

Así,

$$|f| = f^+ + f^-.$$

Si $\bar{h}(\cdot) = h(\cdot) - E[h(Z)]$, *entonces*

$$\begin{aligned} |E[\bar{h}(W)]| &\leq E[|\bar{h}(W)|] \\ &= E[\bar{h}^+(W)] + E[\bar{h}^-(W)] \\ &= E[h^+(W) - h^+(Z)] + E[h^-(W) - h^-(Z)] \end{aligned}$$

Finalmente enunciaremos dos Lemas que serán de gran importancia para la demostración del Teorema :

Lema 2.4.1

Sea $x = x_{\lambda,A}$ *una función definida como*

$$\begin{aligned} x(0) &= 0 \\ x(m+1) &= \lambda^{-(m+1)} e^{-\lambda} m! [P_\lambda(A \cap U_m) - P_\lambda(A) P_\lambda(U_m)]; \quad m \geq 0. \end{aligned}$$

Donde

$A \subset Z^+$.

$U_m = \{0, 1, 2, \ldots, m\}$

$P_\lambda(B) = P(Z \in B)$, *con* $Z \sim Poisson(\lambda)$.

Sea $\Delta f(m) = f(m+1) - f(m)$ *y* $\|\Delta f\| = \sup_{m \in Z^+} |x(m+1) - x(m)|$. *Entonces*

i) $\|x\| \leq \min\{1, 1.25\lambda^{-1/2}\}$.

ii) $\|\Delta x\| \leq \lambda^{-1}(1 - e^{-\lambda}) \leq \min\{1, \lambda^{-1}\}$.

Lema 2.4.2

Con las mismas hipótesis del lema anterior, si $h(w) = 1_{(w=0)} - e^{-\lambda}$, *entonces*

$$\|x\| = \frac{1 - e^{-\lambda}}{\lambda}.$$

La demostraciòn de estos lemas se puede encontrar en el apéndice 1.

2.5 Demostración del teorema principal.

Ahora sí tenemos todos los elementos para entrar de lleno a la demostración del teorema principal de este trabajo.

Sea h dada con $\|h\| = 1$. Sea $\tilde{h}(\cdot) = h(\cdot) - E[h(Z)]$ y $f = S\tilde{h}$.

Notemos que $Tf = \tilde{h}$ y que

$$\begin{aligned} E[Tf(W)] &= E[\tilde{h}(W)] = E[h(W) - E[h(Z)]] \\ &= E[h(W) - h(Z)]. \end{aligned}$$

Ahora:

$$\begin{aligned} E[h(W) - h(Z)] &= E[Tf(W)] = E[Wf(W) - \lambda f(W+1)] \\ &= E\left[\sum_{\alpha\in I}(X_\alpha f(W)) - \sum_{\alpha\in I}(p_\alpha f(W+1))\right] \\ &= \sum_{\alpha\in I} E[X_\alpha f(W) - p_\alpha f(W+1)] \\ &= \sum_{\alpha\in I} E[p_\alpha f(W_\alpha+1) - p_\alpha f(W+1)] \\ &+ \sum_{\alpha\in I} E[X_\alpha f(W_\alpha+1) - p_\alpha f(W_\alpha+1)] \quad \text{Por la observación 2.4.1} \\ &= \sum_{\alpha\in I} E[p_\alpha X_\alpha [f(W_\alpha+1) - f(W_\alpha+2)] \\ &+ \sum_{\alpha\in I} E[X_\alpha f(W_\alpha+1) - p_\alpha f(W_\alpha+1)] \quad \text{Por la observación 2.4.2} \\ &= \sum_{\alpha\in I} E[p_\alpha X_\alpha [f(W_\alpha+1) - f(W_\alpha+2)] \\ &+ \sum_{\alpha\in I} E[(X_\alpha - p_\alpha)\{f(W_\alpha+1) + f(V_\alpha+1) - f(V_\alpha+1)\}] \\ &= \sum_{\alpha\in I} E[p_\alpha X_\alpha [f(W_\alpha+1) - f(W_\alpha+2)] \\ &+ \sum_{\alpha\in I} E[(X_\alpha - p_\alpha)(f(W_\alpha+1) - f(V_\alpha+1))] \\ &+ \sum_{\alpha\in I} E[(X_\alpha - p_\alpha)(f(V_\alpha+1))] \end{aligned}$$

Y usando respectivamente las observaciones 2.4.3, 2.4.4 y 2.4.5 a cada uno de los sumandos en la última igualdad tenemos que

$$\begin{aligned} E[h(W)-h(Z)] &\leq \|\Delta f\| \sum_{\alpha\in I} p_\alpha^2 + \|\Delta f\| \sum_{\alpha\in I} \sum_{\alpha\neq\beta\in B_\alpha} X_\beta + b_3'\|f\| \\ &= \|\Delta f\| (b_1+b_2) + \|f\| b_3' \end{aligned}$$

Ahora utilicemos la observación 2.4.6 para obtener

$$|E[\tilde{h}(W)]| \leq 2[(b_1+b_2)\|\Delta f\| + b_3'\|f\|].$$

Si ahora utilizamos el lema 2.4.1

$$E[\tilde{h}(W)] \leq 2\left[(b_1+b_2)\left(\frac{1-e^{-\lambda}}{\lambda}\right) + b_3' \min\{1, 1.25\lambda^{-1/2}\}\right].$$

Pero $(1-e^{-\lambda})/\lambda \leq 1$ pues $1-\lambda \leq e^{-\lambda} \forall \lambda \in \mathbb{R}$ y $b_3' \leq b_3$, por lo tanto

$$|E[\tilde{h}(W)]| \leq 2(b_1+b_2+b_3).$$

Finalmente, la última parte del teorema se verifica utilizando la equivalencia de la distancia de la variación total dada en la página 47 y el lema 2.4.2.

Capítulo 3

Aplicaciones.

3.1 Un problema de gráficas aleatorias.

Este problema fue estudiado inicialmente por Rinott. La idea principal del problema es tomar los vértices de un cubo en $\mathbb{R}^n$ y asignarle una dirección aleatoria a las aristas. Queremos conocer probabilidades acerca de las direcciones de las flechas que apuntan a tales vértices. En el Apéndice 2 se pueden encontrar las definiciones de teoría de gráficas que se utilizaron en este ejemplo.

Consideremos el cubo $\{0,1\}^n = \{(a_1,\ldots,a_n)|a_i = 0 \text{ ó } 1, \text{ para } i = 1,\ldots,n.\}$ y definamos las aristas de la siguiente manera: Sea e_i el vector canónico en $\mathbb{R}^n$ que tiene 1 en la i-ésima coordenada y 0 en las demás. Sean α y β dos puntos (vértices) en $\{0,1\}^n$. α y β se encuentran conectados si

$$\beta = \alpha + e_i \pmod 2 \quad \text{para alguna } i \in \{1,\ldots,n\}$$

o equivalentemente

$$\alpha = \beta + e_i \pmod 2 \quad \text{para alguna } i \in \{1,\ldots,n\}$$

Convirtamos cada arista en arco asignándole una dirección arbitraria por medio de un volado.

Definición 3.1.1

Sea β un vértice fijo. Diremos que el arco (α,β) apunta hacia adentro (de β) si

comienza en α y termina en β, y apunta hacia afuera si empieza en β y termina en α.

Figura 1: Hacia adentro. Hacia afuera.

Queremos conocer la distribución del número de vértices para los cuales los n arcos que los conectan apuntan hacia adentro.

Sea W el número de vértices para los cuales los n arcos que los conectan apuntan hacia adentro.

Aquí $I = \{0,1\}^n$. Para $\alpha \in I$

$$X_\alpha = \begin{cases} 1 & \text{si los } n \text{ arcos apuntan hacia adentro.} \\ \\ 0 & \text{en otro caso.} \end{cases}$$

Como cada arco apunta hacia adentro con probabilidad 1/2 y hay n arcos, entonces

$$p_\alpha = \frac{1}{2^n}.$$

Para calcular λ, notemos que $|I| = 2^n$, así

$$\lambda = \sum_{\alpha\in I} p_\alpha = \sum_{i=1}^{2^n} \frac{1}{2^n} = 2^n\left(\frac{1}{2^n}\right) = 1.$$

Tomemos $B_\alpha \equiv \{\beta : |\alpha - \beta| = 1\}$, es decir, el conjunto de los vértices adyacentes. Notemos que $|B_\alpha| = n$.

Si α y b son vecinos, entonces $X_\alpha X_\beta = 0$, por lo tanto $E[X_\alpha X_\beta] = 0$. Así

$$\begin{aligned} b_2 &= 0. \\ b_3 &= 0 \quad \text{por independencia.} \\ b_1 &= \sum_{\alpha\in I}\sum_{\beta\in B_\alpha} p_\alpha p_\beta = \sum_{i=1}^{2^n}\sum_{j=1}^{n}\left(\frac{1}{2^n}\right)^2 \\ &= n2^n\left(\frac{1}{2^n}\right)^2 = n2^{-n}. \end{aligned}$$

Finalmente, por el teorema 2.3.1,

$$\|\mathcal{L}(W) - \mathcal{L}(Z)\| \leq 2b_1 = 2n2^{-n}.$$

Lo cual quiere decir que

$$W \xrightarrow{D} Z \qquad \text{cuando} \qquad n \to \infty.$$

Existen otras variantes del problema, aquí daremos una en la cual $\lambda \to \infty$ al mismo tiempo que el método de Chen-Stain funciona, así se puede establecer una aproximación normal.

En el mismo cubo, con las mismas aristas aleatorias, queremos conocer la distribución del número de vértices para los cuales exactamente k arcos apuntan hacia afuera, para $0 \leq k \leq n$.

Sea $W(k,n)$ el número de vértices para los cuales exactamente k arcos apuntan hacia afuera. Así el caso especial $k = 0$ fue el que estudiamos previamente.
Sean $I = \{0,1\}^n$ y para $\alpha \in I$

$$X_\alpha = \begin{cases} 1 & \text{si exáctamente } k \text{ arcos apuntan hacia afuera.} \\ 0 & \text{en otro caso.} \end{cases}$$

Así tenemos que

$$p_\alpha = \frac{\binom{n}{k}}{2^n} \qquad \text{y} \qquad \lambda = \binom{n}{k}.$$

Como antes $B_\alpha \equiv \{\beta : |\alpha - \beta| = 1\}$ y $b_3 = 0$. Ahora:

$$b_1 = \sum_{i=1}^{2^n} \sum_{j=1}^{n} \left(\frac{\binom{n}{k}}{2^n} \right)^2 = n2^{-n} \binom{n}{k}^2 = np_\alpha \lambda.$$

Para $|\alpha - \beta| = 1$, podemos suponer sin pérdida de generalidad que el arco que va de α a β. Por lo tanto

$$\begin{aligned} p_{\alpha\beta} &= E[X_\alpha X_\beta] = P(X_\alpha X_\beta = 1) = \left(\frac{\binom{n-1}{k-1}}{2^n} \right) \left(\frac{\binom{n-1}{k}}{2^n} \right) \\ &\leq p_\alpha^2, \end{aligned}$$

de lo que se concluye que $b_2 \leq b_1$.
Finalmente

$$\|\mathcal{L}(W(k,n)) - \mathcal{L}(Z)\| \leq \frac{2(b_1+b_2)}{\lambda} \leq \frac{4b_1}{\lambda} = 4np_\alpha.$$

Notemos que hay casos para los cuales $b_1 \to \infty$ mientras que $b_1/\lambda \to 0$ y se establece una convergencia Poisson. Notemos también cómo el método de Chen-Stein da lugar a un teorema de límite central:

> Para k, n con $0 < k < n$ y $n2^{-n}\binom{n}{k} \to 0$, la variable aleatoria
>
> $$\frac{W(k,n)) - \binom{n}{k}}{\sqrt{\binom{n}{k}}}$$
>
> converge en distribución a una Normal estándar.

3.2 El problema del cumpleaños.

> "Aunque las coincidencias de nacimientos y muertes que se muestran en estas páginas pueden parecer un poco corrientes, el hecho es que ocurren con regularidad matemática. Los matemáticos conscientes de ésto han hecho, desde hace muchos años, de los nacimientos coincidentes, un juego de salón. Durante una cena con 22 militares, el eminente matemático Warren Weaver hablaba de la probabilidad en favor de dichos aniversarios dobles, y empezó a comparar nacimientos. Llegó hasta el último comensal sin obtener ni una sola coincidencia, pero la camarera, que había escuchado atentamente, anunció de repente que había nacido el mismo día que uno de los invitados." [1]

Si quisieramos hacer un juego parecido con los 60 gobernantes que ha tenido México, desde Agustín de Iturbide hasta Ernesto Zedillo, encontraríamos que al menos 2 nacieron el mismo día: Sebastián Lerdo de Tejada el 24 de abril de 1828 y Manuel

[1] Extracto tomado de [5], pág. 143.

Ávila Camacho, en 1897. Decimos que al menos dos pues, debido a la escasa informacion al respecto, sólamente pudimos averiguar las fechas de nacimiento de 40 gobernantes.

Además, al menos 7 de tales personajes nacieron en marzo, lo cual resulta sorprendente:

- Miguel Barragán, marzo 8. 1789.
- Manuel de la Peña y Peña, marzo 10. 1789.
- Gustavo Díaz Ordaz, marzo 11. 1911.
- Ignacio Comonfort, marzo 12. 1812.
- Benito Juárez, marzo 21. 1806.
- Roque González Garza, marzo 23. 1885.
- Félix Zuloaga, marzo 31. 1813.

Comenzemos a analizar cómo es que el método de Chen-Stein se puede aplicar a este problema de coincidencias en cumpleaños.

En la formulación usual del problema del cumpleaños, suponemos que los cumpleaños de n personas se distribuyen uniforme e independientemente sobre los d días en un año.

Queremos calcular la probabilidad de que k o más personas cumplan años el mismo día, para $k = 2, 3, \ldots$

Comencemos por analizar el caso $k = 2$, esto es, el número de pares de personas que cumplen años el mismo día.

Denotemos por W el número de pares de personas que cumplen años el mismo día. Así $P(W = 0) = 0$ si $n > d$. Supongamos que $n \leq d$.

$$\begin{aligned} P(W = 0) &= \frac{d(d-1)\cdots(d-n+1)}{d^n} = \frac{\binom{d}{n} n!}{d^n} \\ &= \prod_{i=1}^{n-1}\left(1 - \frac{i}{d}\right). \end{aligned}$$

Si ahora estuvieramos interesados en calcular la probabilidad de, digamos, exactamente m coincidencias ($k = m$), o la probabilidad de que al menos tres personas compartan la misma fecha de cumpleaños, o que dos personas cumplieran años la misma semana, o el mismo mes, o probabilidades bajo distribuciones no uniformes en los cumpleaños, entonces los maneras de contar se vuelven mucho menos tratables. Sin embargo, mediante el teorema 2.3.1, se obtienen aproximaciones bastante buenas y fáciles. Además nos da una cota para el error.

Empezaremos por considerar el problema general del cumpleaños cuando éstos son uniformes. Denotemos por $\{1,2,\ldots,n\}$ el grupo de n personas y sea

$$I \equiv \{\alpha \subset \{1,2,\ldots,n\} : |\alpha| = k \leq n\}.$$

Notemos que

$$|I| = \binom{n}{k}.$$

Por ejemplo en el caso anterior $k = 2$ e I es el conjunto de todos los pares de personas entre las cuales podría ocurrir una coincidencia.

Sea X_α el indicador del evento de que las personas indexadas por α compartan la misma fecha de cumpleaños. Ahora, el número total de coincidencias está dado por

$$W = \sum_{\alpha \in I} X_\alpha.$$

Puesto que W es una variable aleatoria Bernoulli, cada una con probabilidad $p_\alpha = d^{1-k}$, parece razonable aproximar W por medio de una variable aleatoria Poisson con media $\lambda = E[W]$:

$$\lambda = E\left[\sum_{\alpha \in I} X_\alpha\right] = \sum_{\alpha \in I} E[X_\alpha] = \binom{n}{k} d^{1-k}$$

y la probabilidad de no coincidencia es aproximadamente

$$P(Z = 0) = e^{-\lambda} = \exp\left\{-\binom{n}{k} d^{1-k}\right\},$$

es decir,

$$P(\text{Ningún día es aniversario de } k \text{ o más personas}) = P(W = 0) \approx P(Z = 0).$$

Sean $\alpha, \beta \in I$. Notemos que si $\alpha \cap \beta = \emptyset$ entonces X_α y X_β son independientes. Esto nos sugiere tomar

$$B_\alpha \equiv \{\beta \in I : \alpha \cap \beta \neq \emptyset\}$$

como nuestro conjunto de dependencia. Al hacer esta selección

$$E\left[|E\left[(X_\alpha - p_\alpha)|\sigma(X_\beta : \beta \notin B_\alpha)\right]|\right] = 0$$

por independencia, así $b_3 = 0$.
Dado que

$$|B_\alpha| = \begin{cases} \binom{n}{k} - \binom{n-k}{k} & \text{si } k \leq n/2. \\ |I| & \text{si } k > n/2. \end{cases}$$

tenemos que

$$\begin{aligned} b_1 &= \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} p_\alpha p_\beta = |I||B_\alpha| p_\alpha^2 \\ &= \frac{\lambda^2 |B_\alpha|}{|I|} < \frac{\lambda^2 k^2}{n} \end{aligned}$$

con igualdad asintótica cuando $n \to \infty$.
En el caso clásico $k = 2$ tenemos que para todo $\alpha \neq \beta$, $p_{\alpha\beta} = p_\alpha p_\beta$. Ahora

$$\begin{aligned} b_2 &= |I|(|B_\alpha| - 1)p_{\alpha\beta} = \frac{b_1(|B_\alpha| - 1)}{|B_\alpha|} \\ &< b_1 \leq \frac{4\lambda^2}{n}, \end{aligned}$$

así

$$\begin{aligned} \|\mathcal{L}(W) - \mathcal{L}(Z)\| &\leq 2\left[\left(\frac{4\lambda^2}{n} + \frac{4\lambda^2}{n}\right)\left(\frac{1 - e^{-\lambda}}{\lambda}\right)\right] \\ &= \frac{16\lambda^2(1 - e^{-\lambda})}{\lambda n}. \end{aligned}$$

Además

$$\begin{aligned} |P(W = 0) - e^{-\lambda}| &\leq (b_1 + b_2)\frac{1 - e^{-\lambda}}{\lambda} \\ &= \frac{1}{d^2}\binom{n}{2}(4n - 7)\frac{1 - e^{-\lambda}}{\lambda}. \end{aligned}$$

En efecto
$$\binom{n-2}{2} = \binom{n}{2}\frac{(n-2)(n-3)}{n(n-1)},$$
por lo tanto
$$|B_\alpha| = \binom{n}{2}\left[1 - \frac{(n-2)(n-3)}{n(n-1)}\right] = \binom{n}{2}\frac{4n-6}{n(n-1)},$$
y entonces
$$2|B_\alpha| - 1 = 4n-7.$$
De aquí que
$$\begin{aligned} b_1 + b_2 &= b_1 + b_1\left(\frac{|B_\alpha|-1}{|B_\alpha|}\right) = b_1\left(1 + \frac{|B_\alpha|-1}{|B_\alpha|}\right) \\ &= \frac{b_1}{|B_\alpha|}(2|B_\alpha|-1) = \binom{n}{2}d^{-2}(4n-7). \end{aligned}$$
Así finalmente,
$$b_1 + b_2 = \frac{\binom{n}{2}}{d^2}(4n-7).$$

Aunque es más dificil calcular exactamente la probabilidad de coincidencia triple en un cumpleaños, podemos aplicar la aproximación Poisson como en el caso clásico. Supongamos que queremos calcular la probabilidad de que en un grupo de 50 personas, tres o más hayan nacido el mismo día del año. Tenemos que $\lambda = \binom{n}{3}/d^2$ y
$$\begin{aligned} P(W \neq 0) &= 1 - P(W=0) \approx 1 - e^{-\lambda} = 1 - 0.863 \\ &= 0.137. \end{aligned}$$
Ahora bién, para calcular una cota para el error, podemos calcular
$$b_1 = |I||B_\alpha|p_\alpha^2 = \binom{n}{3}\left\{\binom{n}{3} - \binom{n}{n-3}\right\}d^{-4}$$
y para α dado, partiendo $B_\alpha - \{\alpha\}$ en aquellos β para los cuales $|\beta \cap \alpha| = 1$ y aquellos para los cuales $|\beta \cap \alpha| = 2$, vemos que
$$b_2 = |I|\left\{3\binom{n-3}{2}d^{-4} + 3(n-3)d^{-3}\right\}.$$

Esto muestra que la aproximación anterior tiene un error no mayor que

$$\frac{(b_1+b_2)(1-e^{-\lambda})}{\lambda} = 0.0597.$$

Así

$$0.803 \le P(W=0) \le 0.923.$$

Para el caso general $k \ge 2$, tenemos

$$b_2 = \sum_{\alpha\in I}\sum_{\alpha\neq\beta\in B_\alpha} p_{\alpha\beta}.$$

Para j fija tal que $j = |\alpha\cap\beta|$

$$p_{\alpha\beta} = \frac{1}{d^{2k-j-1}} = d^{1+j-2k},$$

pues $|\alpha\cup\beta| = 2k-j$. Ahora

$$|B_\alpha| = \binom{k}{1}\binom{n-k}{k-1} + \binom{k}{2}\binom{n-k}{k-2} + \cdots + \binom{k}{K+1}\binom{n-k}{1}.$$

Donde el i-ésimo término de la suma se puede interpretar como sigue:

> de los k elementos de α tomamos i que serán los elementos de la intersección; de los $n-k$ restantes tomamos $k-i$, para que sólo se intersecten en i elementos.

Así

$$\begin{aligned} b_2 &= \sum_{\alpha\in I}\sum_{j=1}^{k-1}\binom{k}{j}\binom{n-k}{k-j}d^{1+j-2k} \\ &= \sum_{j=1}^{k-1}\binom{n}{k}\binom{k}{j}\binom{n-k}{k-j}d^{1+j-2k} \end{aligned}$$

Ahora tomemos $n,d\to\infty$ de tal manera que la razón $\lambda/1$ se encuentre acotada lejos de cero y de infinito, lo cual denotaremos por $\lambda \asymp 1$. Entonces $n^k \asymp d^{k-1}$, $b_1 \asymp |B_\alpha|/|I| \asymp n^{-1}$ y $b_2 \asymp n^{1+k}d^{-k} \asymp n/d \asymp n^{-1/(k-1)}$.
En efecto:

i)$n^k \asymp d^{k-1}$.
Utilizaremos la aproximación de Stirling:

$$n! \approx \sqrt{2\pi} n^{n+(1/2)} e^{-n}.$$

$$\begin{aligned} \frac{n! d^{1-k}}{k!(n-k)!} &\approx \frac{\sqrt{2\pi} n^{n+(1/2)} e^{-n}}{k! \sqrt{2\pi} (n-k)^{n-k+(1/2)} e^{n-k} d^{k-1}} \\ &= \frac{1}{k! e^k} \left(\frac{n}{n-k}\right)^{n+(1/2)} \left(\frac{n-k}{n}\right)^k \left(\frac{n^k}{d^{k-1}}\right) \end{aligned}$$

y considerando que $1/(k!e^k)$ es constante con respecto a n y a d, y que $(n/(n-k))^{n+1/2}$ y $((n-k)/n)^k$ tienden a 1 caundo $n \to \infty$, tenemos el resultado.

ii)$b_1 \asymp |B_\alpha|/|I|$
Solamente hace falta hacer notar que

$$\frac{b_1}{|B_\alpha|/|I|} = \lambda^2.$$

iii)$|B_\alpha|/|I| \asymp n^{-1}$
Claro, pues $|B_\alpha|/|I| \approx k^2/n$.
De manera análoga se muestran las demás relaciones. Así para $k \geq 3$, tenemos que $\|\mathcal{L}(W) - \mathcal{L}(Z)\| = O(n^{-1/(k-1)})$ con b_2 haciendo la contribución principal, es decir, la distancia de la variación total decae a una tasa no más lenta que $\|\mathcal{L}(W) - \mathcal{L}(Z)\| = O(n^{-1/(k-1)})$.

Antes de terminar, vale la pena hacer notar que las cotas pueden ser mejoradas si cambiamos la definición de B_α por una menos natural:

$$B_\alpha \equiv \{\beta \in I : (\alpha - min(\alpha)) \cap \beta \neq \emptyset\}.$$

Así

$$|B_\alpha| = \binom{n}{k-1} - \binom{n-(k-1)}{k-1}.$$

3.3 Rachas en lanzamientos de monedas.

En esta seccion analizaremos la longitud de la racha más larga en lanzamientos de monedas. Consideremos una sucesión muy grande de lanzamientos independientes de una moneda con probabilidad p de que caiga sol, con $0 < p < 1$. Sin importar cual sea el valor de p, habrá algunos lapsos donde sólo caiga sol.[3] Denotemos por R_n la longitud más grande en rachas que empieza en los primeros n ensayos. Antes de empezar a analizar la distribución de R_n, notemos que, para una longitud de prueba t apropiadamente elegida, uno puede ver, con una probabilidad muy pequeñá, una racha de t soles que empiezan en una posición dada α. Como el número de tales posiciones donde tales rachas pueden ocurrir es muy grande, y la probabilidad de ocurrencia es muy pequeña, parece válido hacer una aproximación Poisson.

Sean $C_1, C_2, \ldots$ variables aleatorias independientes Bernoulli con probabilidad de éxito $p = P(C_i = 1) = 1 - P(C_i = 0)$.

Sea $I = \{1, 2, \ldots, n\}$ los elementos del conjunto de índices que denotarán los lugares donde las rachas pueden empezar. Sea $t \in \mathbb{N}$ una longitud de prueba fija.

Una racha de soles de longitud t o más empieza en la posición α si y sólo si la variable indicadora

$$Y_\alpha = \prod_{i=\alpha}^{\alpha+t-1} C_i$$

toma el valor de uno. Para contar solamente la primera racha en un grupo tomamos $X_1 = Y_1$ y

$$X_\alpha = (1 - C_{\alpha-1})Y_\alpha, \qquad \alpha = 2, 3, \ldots, n.$$

Para $a = 2, 3, \ldots, n$, X_a tomará el valor de uno si y solamente si una racha de longitud t empieza en la posición α y cayó águila en la posición $\alpha - 1$.

Si no hubieramos elegudo α de esta manera y solamente hubieramos tomado $X_\alpha = Y_\alpha$, hubieramos tenido que b_2 no tiende a cero y, de hecho, no hubiera sido válida una aproximación Poisson.

[3]Ver [8] $VIII.3$ Lema 2.

Denotemos por W el número total de grupos de rachas de t o más soles, entonces

$$W = \sum_{\alpha \in I} X_\alpha.$$

La aproximación Poisson heurística dice que deberíamos poder aproximar la distribución de W por una variable aleatoria Poisson con media λ, donde

$$\begin{aligned}
\lambda &= \lambda_n(t) = E[W] = E[X_1] + \sum_{j=2}^{n} E[X_j] \\
&= p^t + (n-1)p^t(1-p) \\
&= p^t[1 + (n-1)(1-p)]
\end{aligned}$$

Dado que los eventos $\{R_n < t\}$ y $\{W = 0\}$ son iguales, la función de distribución de R_n puede ser aproximada como

$$P(R_n < t) = P(W = 0) = e^{-\lambda_n(t)}.$$

Sea $B_\alpha = \{\beta \in I : |\beta - \alpha| < t\}$. Como X_α es independiente de $\sigma(X_\beta : \beta \in B_\alpha)$ entonces $b_3 = 0$. Mas aún, si $1 < |\beta - \alpha| < t$, no puede ser que X_α y X_β tomen el valor de uno al mismo tiempo si $\beta \in B_\alpha$, $\alpha \neq \beta$, pues ya habría habido una racha antes de α y queremos contar sólo la primera racha. Por lo que $b_2 = 0$.

Para calcular $b_1 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} p_\alpha p_\beta$, dividimos la suma en dos partes, dependiendo de si esta o no p_1, y entonces, como $|B_\alpha| \leq 2t - 2$

$$\begin{aligned}
b_1 &= \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} p_\alpha p_\beta = \sum_{i=1}^{n} \sum_{j \in B_i} p_i p_j \\
&= 2p_1 \sum_{j \in B_1} p_j + \sum_{i=2}^{n} \sum_{j \in B_i} p_i p_j \\
&\leq 2p^t \sum_{j \in I} p_j + (n-1)(2t-2)p^{2t}(1-p)^2 \\
&= 2\lambda p^t + \frac{(n-1)^2(2t-2)p^{2t}(1-p)^2}{n-1} \\
&< 2\lambda p^t + \frac{(2t-2)\lambda^2}{n-1}.
\end{aligned}$$

Por lo tanto

$$b_1 < 2\lambda p^t + \frac{(2t-2)\lambda^2}{n-1}.$$

Así tenemos que

$$|P(R_n < t) - e^{-\lambda_n(t)}| \; < \; b_1 \min\{1, \lambda^{-1}\}.$$

Capítulo 4

Conclusiones

Desde 1971, año en que el Dr. Charles Stein publicó su artículo de aproximación Normal en distribución para una suma de variables aleatorias dependientes, se han escrito numerosas publicaciones que tratan de seguir el camino que éste había iniciado. El primero en hacerlo fué Louis Chen, quien fuera alumno de Doctorado del Dr. Stein. Trazó otro camino en cuanto a aproximaciones, pero esta vez, para distribuciones discretas (de ahí el nombre de método de Chen-Stein).

Después de él, otros autores como Arratia, Barbour, Goldstein, Gordon, Eagleson y otros, han escrito artículos acerca de estos temas. No se si se deba a la tal vez "reciente" aparición de éstos artículos, pero la verdad es que aún no son muy conocidos y, a mi parecer, constituyen una gran herramienta cuando se estan haciendo cálculos de probabilidades.

La gran importancia de este método, se basa en que no solamente nos dice que podemos trabajar con variables aleatorias Poisson si los parámetros que se definieron en el capítulo 2 son pequeños, sino que además nos dá una cota para el error de la aproximación. Y no sólo eso, sino que sólamente estamos haciendo el cálculo de los dos primeros momentos de la distribición, lo cuál no representa demasiado esfuerzo.

Si bien es cierto que en ciertas ocaciones los parámetros no son fáciles de calcular con exactitud, la mayoría de las veces se pueden encontrar cotas superiores para estas cantidades.

En el presente trabajo, se intentó presentar el método de Chen-Stein de una manera que mostrara la fácilidad con la que éste se aplica. En los ejemplos presentados, intentamos mostrar la manera en la que se aplica el método y cómo se pueden salvar algunos problemas que pudieran encontrarse a lo largo del desarrollo de la aplicación del mismo. Por ejemplo, en el problema de las gráficas aleatorias se pudieron encontar con relativa facilidad los parámetros e incluso se pudo extender la aproximación a una Normal estándar bajo ciertas condiciones. En las dos últimas aplicaciones se encontraron algunas dificultades para calcular tales parámetros, pero se pudieron acotar de una manera muy interesante.

Como dato adicional cabe hacer notar que en el resultado principal de éste trabajo, las cotas presentadas fueron mejoradas: en el artículo original, el factor para β_3' es

$$\min\{1, 1.4\lambda^{-1/2}\}$$

y en la tesis aparece como

$$\min\{1, 1.25\lambda^{-1/2}\}.$$

Esto se debió a la reestructuración que se hizo en la demostración de los lemas 2.4.1 y 2.4.2.

Cabe señalar que el resultado principal que se dá aquí no es el único, de hecho, existen extenciones para el caso de Procesos Estocásticos que se presentan en el mismo artículo. Los ejemplos que se dan de ninguna menera son los únicos ni en las únicas ramas en las que se pueden aplicar. Existe una gran variedad de ejemplos que se pueden encontrar en [1].

Finalmente vale la pena mencionar que este tipo de aproximaciones realmente tienen aplicaciones reales. Por ejemplo, en los Estados Unidos se está haciendo uso de estos resultados en juicios de herencias utilizando las cadenas de DNA de los individuos en disputa.

Apéndice 1.

En este apartado se presenta la demostración de los lemas 2.4.1 y 2.4.2. Antes de entrar de lleno a las demostraciones, haremos algunas observaciones que nos ayudaran en tales pruebas.

Sean

$$\begin{aligned} P_\lambda(A) &= P(Z \in A), \qquad \text{con } Z \sim \mathit{Poisson}(\lambda). \\ A &\subset \mathcal{Z}^+. \\ U_m &= \{0, 1, \ldots, m\}. \\ B^- &= \text{El complemento de } B. \end{aligned}$$

Observación 1

$$\left| P_\lambda(A \cap U_m) P_\lambda(U_m^-) - P_\lambda(A \cap U_m^-) P_\lambda(U_m) \right| \leq P_\lambda(U_m^-) P_\lambda(U_m).$$

Demostración:

$$|a - b| = \begin{cases} a - b & \text{si } b \leq a \\ b - a & \text{si } b > a \end{cases}$$

Por lo tanto

$$|a - b| = \begin{cases} a & \text{si } b \leq a \\ b & \text{si } b > a \end{cases}$$

De aquí que si

$$\begin{aligned} a &= P_\lambda(A \cap U_m) P_\lambda(U_m^-) \\ b &= P_\lambda(A \cap U_m^-) P_\lambda(U_m) \end{aligned}$$

Entonces
$$|a-b| \leq P_\lambda(A\cap U_m)P_\lambda(U_m^-) \;\leq\; P_\lambda(U_m)P_\lambda(U_m^-)$$
y también
$$|a-b| \leq P_\lambda(A\cap U_m^-)P_\lambda(U_m) \;\leq\; P_\lambda(U_m^-)P_\lambda(U_m).$$
Por lo tanto
$$\left|P_\lambda(A\cap U_m)P_\lambda(U_m^-) - P_\lambda(A\cap U_m^-)P_\lambda(U_m)\right| \leq P_\lambda(U_m^-)P_\lambda(U_m).$$

□

Observación 2

Sean
$$f(\lambda) = \frac{1+\lambda}{\lambda^2}; \qquad g(\lambda) = 1 - e^{-\lambda}(1+\lambda); \quad \lambda \geq 0.$$
entonces, si $12/7 \leq \lambda \leq 2$ *se tiene que*
$$f(\lambda)g(\lambda) \leq \left(\frac{133}{144}\right)\left(1-3e^{-2}\right).$$

Demostración:

Si derivamos f y g podemos checar que para $\lambda > 0$, f es una función decreciente y g creciente, por lo tanto, sus valores máximos se alcanzan respectivamente en $12/7$ y en 2. Así
$$f(\lambda)g(\lambda) \leq f(12/7)g(2) \leq \left(\frac{133}{144}\right)\left(1-3e^{-2}\right).$$

□

Observación 3

Sea $a \in \mathbb{R}$, *entonces* $a(1-a) \leq 1/4$.

Demostración:

$$a(1-a) \leq \frac{1}{4} \iff \left(a - \frac{1}{2}\right)^2 \geq 0.$$

□

Observación 4

$$\left(\frac{m+1}{2}\right)+\left(m+\frac{1}{2}\right)\ln(m)\leq m^2.$$

Demostración:

La demostración se hará por inducción.

Para $m=1$ se da la igualdad. Antes de continuar debemos notar dos cosas:

i) $\ln(m+1)\leq\ln(m)+1 \Longleftrightarrow m+1\leq me \Longleftrightarrow 1+\frac{1}{m}\leq e$ lo cual es cierto para $m\geq 1$.

ii) $\ln(m+1)\leq m \Longleftrightarrow m+1\leq e^m$ lo cual es cierto para $m\geq 1$.

Teniendo en cuenta lo anterior supongamos que nuestro resultado es válido para m, es decir,

$$\left(\frac{m+1}{2}\right)+\left(m+\frac{1}{2}\right)\ln(m)\leq m^2.$$

Demostraremos que es válido para $m+1$.

$$\begin{aligned}
\left(\frac{m+1+1}{2}\right)\Big(m &+ 1+\frac{1}{2}\Big)\ln(m+1)\\
&= \frac{m+1}{2}+\left(m+\frac{1}{2}\right)\ln(m+1)+\frac{1}{2}+\ln(m+1)\\
&\leq \frac{m+1}{2}+\left(m+\frac{1}{2}\right)\ln(m)+\left[\left(m+\frac{1}{2}\right)\ln(m+1)\frac{1}{2}\right] \text{Por } i)\\
&\leq m^2+[m+1+\ln(m+1)] \text{ Por hip. de induc.}\\
&= m^2+m+1+\ln(m+1)\\
&\leq m^2+m+1+m=(m+1)^2 \text{Por } ii)
\end{aligned}$$

□

Observación 5

$$\left(\frac{m}{\lambda}\right)^{m+1/2}e^{\lambda-m}\leq\exp\left(\frac{m-\lambda}{\lambda}\right)\left(m-\lambda+\frac{1}{2}\right).$$

Demostración:

i) Sea $\lambda\in(0,1]$

$$\left(\frac{m}{\lambda}\right)^{m+1/2}e^{\lambda-m}\leq\exp\left\{\left(\frac{m-\lambda}{\lambda}\right)\left(m-\lambda+\frac{1}{2}\right)\right\} \Longleftrightarrow$$

$$1 \le \exp\left(\frac{m^2}{\lambda} - m + \frac{m}{2\lambda} - \frac{1}{2}\right)\left(\frac{\lambda}{m}\right)^{m+1/2} \iff$$
$$1 \le \left(\frac{\lambda}{m}\right)^{m+1/2} \exp\left(\frac{m^2}{\lambda} + \frac{1}{2}\frac{m}{\lambda} - \frac{1}{2}\right).$$

$$\begin{aligned}\text{Sea } f(\lambda) &= \left(\frac{\lambda}{m}\right)^{m+1/2} \exp\left(\frac{m^2}{\lambda} + \frac{1}{2}\frac{m}{\lambda} - m - \frac{1}{2}\right)\\ &= \exp\left\{\frac{m^2}{\lambda} + \frac{1}{2}\frac{m}{\lambda} - m - \frac{1}{2} + \left(m + \frac{1}{2}\right)\ln\left(\frac{\lambda}{m}\right)\right\}.\end{aligned}$$

Por lo tanto f es una función monótona y cumple que

$$\lim_{\lambda\to 0} f(\lambda) = \infty \quad \text{y} \quad f(1) = \left(\frac{1}{m}\right)^{m+1/2} \exp\left(m^2 - \frac{m}{2} - \frac{1}{2}\right).$$

Para concluir esta primera parte basta notar que

$$\begin{aligned}1 \le f(1) \iff 1 &\le \left(\frac{\lambda}{m}\right)^{m+1/2} \exp\left(m^2 - \frac{m}{2} - \frac{1}{2}\right)\\ \iff & \left(m + \frac{1}{2}\right)\ln(m) \le m^2 - \frac{m}{2} - \frac{1}{2}\\ \iff & \left(\frac{m+1}{2}\right) + \left(m + \frac{1}{2}\right)\ln(m) \le m^2.\end{aligned}$$

Lo cual es cierto por la observación 4.

ii) Sea $\lambda \in [1, \infty)$

$$\begin{aligned}&\left(\frac{m}{\lambda}\right)^{m+1/2} e^{\lambda - m + 1/12} \le \exp\left\{\left(\frac{m-\lambda}{\lambda}\right)\left(m - \lambda + \frac{1}{2} + \frac{1}{12}\right)\right\}\\ \iff & \left(\frac{m}{\lambda}\right)^{m+1/2} \exp\left(-\frac{m}{2} + m - \frac{1}{2}\frac{m}{\lambda} + \frac{1}{2}\right) \le 1\\ \iff & f(1) \le 1, \text{ pues } \lim_{\lambda\to 0} f(\lambda) = 0 \text{ y por monotonía de la función exponencial}\\ \iff & \left(m + \frac{1}{2}\right)\ln(m) - m^2 + \frac{m}{2} + \frac{1}{2} \le 0\\ \iff & \left(\frac{m+1}{2}\right) + \left(m + \frac{1}{2}\right)\ln(m) \le m^2\end{aligned}$$

Lo cual es cierto otra vez por la observación 4.

Observación 6

$$P(Z \leq j) = 1 - \int_0^\lambda e^{-v}\frac{v^j}{j!}dv.$$

Demostración:
Haremos la demostración por inducción.
Para $j = 0$

$$\begin{aligned} P(Z = 0) &= 1 - \int_0^\lambda e^{-v}dv = 1 + \left(e^{-v}\Big|_0^\lambda\right) \\ &= e^{-\lambda} \end{aligned}$$

Supongamos que la afirmación es válida para $j-1$, por lo tanto

$$P(Z \leq j-1) = 1 - \int_0^\lambda e^{-v}\frac{v^{j-1}}{(j-1)!}dv.$$

Y así

$$\begin{aligned} P(Z \leq j) &= P(Z \leq j-1) + P(Z = j) \\ &= 1 - \int_0^\lambda e^{-v}\frac{v^{j-1}}{(j-1)!}dv + \frac{e^{-\lambda}\lambda^j}{j!}. \end{aligned}$$

Pero

$$\int_0^\lambda e^{-v}\frac{v^j}{j!}dv = -\left(\frac{e^{-\lambda}\lambda^j}{j!}\right) + \int_0^\lambda e^{-v}\frac{v^{j-1}}{(j-1)!}dv$$

Por lo tanto

$$\frac{e^{-\lambda}\lambda^j}{j!} = \int_0^\lambda e^{-v}\frac{v^{j-1}}{(j-1)!}dv - \int_0^\lambda e^{-v}\frac{v^j}{j!}dv.$$

Y, finalmente

$$P(Z \leq j) = 1 - \int_0^\lambda e^{-v}\frac{v^j}{j!}dv.$$

□

A continuación daremos los lemas que serán muy importantes en la demostración de los teoremas principales.
Definamos la función $x = x_{\lambda,A}$ de la siguiente manera:

$$\begin{aligned} x(0) &= 0 \\ x(m+1) &= \lambda^{-(m+1)}e^\lambda m![P_\lambda(A \cap U_m) - P_\lambda(A)P_\lambda(U_m)]; \qquad m \geq 0. \end{aligned}$$

Donde

$$\begin{aligned} A &\subset \mathbb{Z}^+ = \{0,1,2,\ldots,\} \\ U_m &= \{0,1,2,\ldots,m\} \\ P_\lambda(B) &= P(Z \in B), \qquad \text{Con } Z \sim Poisson(\lambda). \end{aligned}$$

Sea $\Delta x = x(m+1) - x(m)$ y $\|\Delta x\| = \sup_{m\in\mathbb{Z}^+} |x(m+1) - x(m)|$.

Lema 1 *Para x definida arriba*

$$\begin{aligned} &i\ )\ \ \|x\| \le \min\{1, 1.25\lambda^{-1/2}\}. \\ &ii\ )\ \ \|\Delta x\| \le \lambda^{-1}\left(1 - e^{-\lambda}\right) \le \min\{1, \lambda^{-1}\}. \end{aligned}$$

Demostración:

Sea B^- el complemento de B. Notemos que

$$\begin{aligned} & P_\lambda(A \cap U_m) - P_\lambda(U_m) P_\lambda(A) \\ = \ & P_\lambda(A \cap U_m) - P_\lambda(U_m)\left[P_\lambda(A \cap U_m) + P_\lambda\left(A \cap U_m^-\right)\right] \\ = \ & P_\lambda(A \cap U_m) - [1 - P_\lambda(U_m)] - P_\lambda\left(A \cap U_m^-\right) P_\lambda(U_m) \\ = \ & P_\lambda(A \cap U_m) P_\lambda\left(U_m^-\right) - P_\lambda\left(A \cap U_m^-\right) P_\lambda(U_m). \end{aligned}$$

Así

$$x(m+1) = \lambda^{-(m+1)} e^\lambda m! [P_\lambda(A \cap U_m) P_\lambda\left(U_m^-\right) - P_\lambda\left(A \cap U_m^-\right) P_\lambda(U_m)].$$

Dado que $0 \le P_\lambda(A \bigcap B) P_\lambda(B) \le P_\lambda(B) P_\lambda(B^-)$ para cualquier $A \subset \mathbb{Z}^+$ y en virtud de la observación 1

$$|x(m+1)| \le m! e^\lambda \lambda^{-(m+1)} P_\lambda(U_m) P_\lambda\left(U_m^-\right). \tag{7}$$

Para $m \le \lambda$

$$\begin{aligned} |x(m+1)| &\le m! e^\lambda \lambda^{-(m+1)} P_\lambda(U_m) \\ &= m! e^\lambda \lambda^{-(m+1)} \sum_{j=0}^{m} e^{-\lambda} \frac{\lambda^j}{j!} \\ &= \frac{1}{\lambda} \sum_{j=0}^{m} \frac{(\lambda^{j-m}) m!}{j!} = \frac{1}{\lambda} \sum_{j=0}^{m} \frac{\lambda^{-j} m!}{(m-j)!} \end{aligned}$$

Donde la última igualdad se verifica al escribir los términos en orden inverso. Ahora

$$\frac{\lambda^{-j}m!}{(m-j)!} \leq \left(\frac{m}{\lambda}\right)^j \Longleftrightarrow \frac{m!}{(m-j)!} \leq m^j \Longleftrightarrow m(m-1)\cdots(m-(j-1)) \leq m^j.$$

Lo cual es cierto. Así

$$\sum_{j=0}^{m} \frac{\lambda^{-j}m!}{(m-j)!} \leq \sum_{j=0}^{m} \left(\frac{m}{\lambda}\right)^j,$$

por lo tanto

$$\begin{aligned} |x(m+1)| &\leq \frac{1}{\lambda}\sum_{j=0}^{m}\left(\frac{m}{\lambda}\right)^j \leq \frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{m}{\lambda}\right)^j \\ &= \frac{1}{\lambda}\left(\frac{\lambda}{\lambda-m}\right) = (\lambda-m)^{-1}, \end{aligned}$$

es decir

$$|x(m+1)| \leq (\lambda-m)^{-1} \qquad \text{si } m \leq \lambda \tag{8}$$

De manera análoga para $m \geq \lambda - 3$

$$\begin{aligned} |x(m+1)| &\leq m!e^{\lambda}\lambda^{-(m+1)}P_\lambda(U_m^-) \\ &= \sum_{j=0}^{\infty} \frac{\lambda^j m!}{(m+1+j)!} \\ &\leq \frac{1}{m+1}\left[1 + \frac{\lambda}{m+2}\sum_{j=0}^{\infty}\left(\frac{\lambda}{m+3}\right)^j\right] \\ &= \frac{(m+2)(m+3)+\lambda}{(m+1)(m+2)(m+3+\lambda)} \end{aligned} \tag{9}$$

De la igualdad(8) :

$$|x(m+1)| \leq 1 \qquad \text{siempre que } m \leq \lambda - 1.$$

De la ecuación (9) :
si $m = 1$ y $\lambda \leq 12/7$ entonces $|x(m+1)| \leq 1$.

Si $m \geq 2$ y $\lambda - 1 \leq m$ entonces

$$\frac{(m+2)(m+3)+\lambda}{(m+1)(m+2)(m+3-\lambda)} \leq \frac{(m+2)(m+3)+(m+1)}{(m+1)(m+2)(m+3)-\lambda)},$$

pero $m+3-\lambda \geq 2$, por lo tanto

$$\begin{aligned}
\frac{(m+2)(m+3)+(m+1)}{(m+1)(m+2)(m+3-\lambda)} &\leq \frac{(m+2)(m+3)+(m+1)}{2(m+1)(m+2)}\\
&= \frac{1}{2}\left[\frac{m+3}{m+1}+\frac{1}{m+2}\right]\\
&= \frac{1}{2}\left[1+\frac{2}{m+1}+\frac{1}{m+2}\right]\\
&\leq \frac{1}{2}\left[1+\frac{2}{3}+\frac{1}{4}\right]\\
&= \frac{23}{24}<1.
\end{aligned}$$

Por lo tanto, hemos probado que

$$\frac{(m+2)(m+3)+\lambda}{(m+1)(m+2)(m+3-\lambda)}<1.$$

Aún nos falta considerar el caso en el que $m=0$ con $0 \leq \lambda < 1$ y $m=1$ con $12/7 \leq \lambda \leq 2$.
Sin embargo, se sigue de (7) que para $m=0$,

$$|x(1)| \leq \lambda^{-1}\left(1-e^{-\lambda}\right) \leq 1 \qquad \forall \lambda \in \mathbb{R}.$$

Para $m=1$ con $12/7 \leq \lambda \leq 2$, y utilizando la observación 2 :

$$\begin{aligned}
|x(2)| &\leq \left[\lambda^{-2}(1+\lambda)\right]\left[1-e^{-\lambda}(1+\lambda)\right]\\
&\leq \left(\frac{133}{144}\right)\left(1-3e^{-2}\right)<1.
\end{aligned}$$

Notemos que, de esta manera, hemos cubierto todos los casos. Así finalmente

$$\|x\| \leq 1.$$

Con lo cual queda demostrada la primera parte del primer inciso del lema.
Regresando a (7), utilizando la observación 3 y la aproximación de Stirling:

$$m! \leq \sqrt{2\pi}\, m^{m+1/2} \exp\left(\lambda - m + \frac{1}{12m}\right)$$

iii) Para $\lambda - (4/5)\lambda^{1/2} \leq m \leq \lambda + (4/5)\lambda^{1/2} - 1$ definimos

$$\begin{aligned}\psi(m) &= (m-\lambda)\left(m-\lambda+\frac{1}{2}\right)\\ &= \left[m-\left(\lambda-\frac{1}{4}\right)\right]^2-\frac{1}{4}.\end{aligned}$$

Es decir, ψ es una parábola que se abre hacia arriba con vértice en $(\lambda - 1/4, -1/4)$ e intersección con el eje de las absisas en $\lambda - 3/4$ y en $\lambda + 1/4$.
Notemos que

$$\lambda - \frac{1}{4} > \lambda + \frac{4}{5}\lambda^{1/2} \Longleftrightarrow \left(\frac{5}{16}\right)^2 < \lambda$$

y que

$$\lambda - \frac{1}{4} < \lambda + \frac{4}{5}\lambda^{1/2} - 1 \Longleftrightarrow \left(\frac{15}{16}\right)^2 < \lambda.$$

Lo cual se cumple pues $\lambda > (1.25)^2$. Así

$$\lambda - \frac{4}{5}\lambda^{1/2} < \lambda - \frac{1}{4} < \lambda + \frac{4}{5}\lambda^{1/2} - 1.$$

Lo cual quiere decir que el máximo de la función ψ se encuentra en el extremo del intervalo que se encuentre más alejado de $\lambda - 1/4$.
Afirmación :
El máximo de ψ se alcanza en $m = \lambda - (4/5)\lambda^{1/2}$.
Demostración:

$$\begin{aligned}&\left(\lambda+\frac{4}{5}\lambda^{1/2}-1\right)-\left(\lambda-\frac{1}{4}\right) < \left(\lambda-\frac{1}{4}\right)-\left(\lambda-\frac{4}{5}\lambda^{1/2}\right)\\ \Longleftrightarrow\ & \frac{4}{5}\lambda^{1/2}-\frac{3}{4} < \frac{4}{5}\lambda^{1/2}-\frac{1}{4}\\ \Longleftrightarrow\ & -\frac{3}{4} < -\frac{1}{4}.\end{aligned}$$

Debido a lo anterior:

$$\begin{aligned}\lambda^{-1}(m-\lambda)\left(m-\lambda+\frac{1}{2}\right) &\leq \frac{1}{\lambda}\left(\lambda-\frac{4}{5}\lambda^{1/2}-\lambda\right)\left(-\frac{4}{5}\lambda^{1/2}-\lambda+\frac{1}{2}\right)\\ &= \frac{1}{\lambda}\left(-\frac{4}{5}\lambda^{1/2}\right)\left(-\frac{4}{5}\lambda^{1/2}+\frac{1}{2}\right)\\ &= \frac{16}{25}-\frac{2}{5}\lambda^{-1/2} \leq \frac{16}{25}.\end{aligned}$$

se sigue que

$$\begin{aligned}|x(m+1)| &\leq \frac{1}{4}\sqrt{2\pi}\lambda^{-1/2}\left(\frac{m}{\lambda}\right)^{m+1/2}\exp\left(\lambda-m+\frac{1}{12m}\right)\\ &\leq \frac{1}{4}\sqrt{2\pi}\lambda^{-1/2}\exp\left(\lambda^{-1}(m-\lambda)\left(m-\lambda+\frac{1}{2}\right)+\frac{1}{12}\right) \qquad m\geq 1.\end{aligned}$$

Donde la última desigualdad se obtiene de la observación 5.

Como antes

$$|x(1)| \leq \lambda^{-1}\left(1-e^{-\lambda}\right) \leq \lambda^{-1}.$$

Ahora tenemos tres casos por analizar

Caso 1:

Si $0 < \lambda \leq (1.25)^2$ entonces $\min\{1, 1.25\lambda^{-1/2}\} = 1$ y por lo tanto $\|x\| \leq 1$ segun lo demostrado anteriormente.

Caso 2:

Si $\lambda > (1.25)^2$ tenemos tres posibilidades:

i)$0 \leq m \leq \lambda - (4/5)\lambda^{1/2}$. En este caso podemos usar 8 para obtener que $\lambda - m \geq (4/5)\lambda^{1/2}$ y por lo tanto $1/(\lambda-m) \leq (5/4)\lambda^{-1/2}$, así $\|x\| \leq 1/(\lambda-m) \leq 1.24\lambda^{-1/2}$.

ii) Notemos que de 9 :

$$\begin{aligned}|x(m+1) &\leq m!e^{\lambda}\lambda^{-(m+1)}P_{\lambda}\left(U_m^-\right)\\ &= \sum_{j=0}^{\infty}\frac{\lambda^j m!}{(m+1+j)!}\\ &\leq \frac{1}{m+1}\sum_{j=0}^{\infty}\left(\frac{\lambda}{m+1}\right)^j \text{(si } m > \lambda - 1)\\ &= \frac{1}{m-\lambda+1}.\end{aligned}$$

Si $m > \lambda + (4/5)\lambda^{1/2} - 1$

$$|x(m+1)| \leq \frac{1}{m-\lambda+1} \leq 1.25\lambda^{-1/2}.$$

pues $m - \lambda + 1 \geq (4/5)\lambda^{1/2}$ si y sólo si $1/(m-\lambda+1) \leq 1.25\lambda^{-1/2}$. Por lo tanto

$$\|x\| \leq 1.25\lambda^{-1/2}.$$

De aquí que:

$$
\begin{aligned}
& \frac{\lambda^{-1/2}}{4} \exp\left\{\lambda^{-1}(m-\lambda)\left(m-\lambda+\frac{1}{2}\right)+\frac{1}{12m}\right\} \\
\leq\ & \frac{\lambda^{-1/2}}{4} \exp\left\{\frac{16}{25}+\frac{1}{24}\right\} \qquad (\text{si } m \geq 2) \\
\approx\ & 1.239\lambda^{-1/2} \\
<\ & 1.25\lambda^{-1/2}.
\end{aligned}
$$

Por lo tanto

$$\|x\| \leq 1.25\lambda^{1/2}.$$

Lo cual concluye la primera parte del lema, es decir,

$$\|x\| \leq \min\{1, 1.25\lambda^{-1/2}\}$$

Para la demostración del inciso *ii*) del lema, sea $x_j = x_{\lambda,\{j\}}$ y notemos que

$$x_{\lambda,A} = \sum_{j\in A} x_j.$$

De esta definición,
a) Para $j \leq m$

$$
\begin{aligned}
x_j(m+1) &= \lambda^{-m-1}m!e^{\lambda}\left[P_\lambda(\{j\}) - P_\lambda(\{j\})P_\lambda(U_m)\right] \\
&= \lambda^{-m-1}m!e^{\lambda}P_\lambda(\{j\})\left[1 - P_\lambda(U_m)\right] \\
&= \lambda^{-m-1}m!\frac{\lambda^j}{j!}P_\lambda(U_m^-) \\
&= \frac{\lambda^{j-1}}{j!}m!\lambda^{-m}P_\lambda(U_m^-)
\end{aligned}
$$

b) Para $j > m$

$$
\begin{aligned}
x_j(m+1) &= \lambda^{-m-1}m!e^{\lambda}\left[-\left(\frac{e^{-\lambda}\lambda^j}{j!}\right)P_\lambda(U_m)\right] \\
&= -\left(\frac{\lambda^{j-1}}{j!}\right)m!\lambda^{-m}P_\lambda(U_m)
\end{aligned}
$$

Así

$$x_j(m+1) = \begin{cases} \left(\frac{\lambda^{j-1}}{j!}\right) m!\lambda^{-m} P_\lambda(U_m^-) & \text{si } m \geq j \\ -\left(\frac{\lambda^{j-1}}{j!}\right) m!\lambda^{-m} P_\lambda(U_m) & \text{si } m < j. \end{cases}$$

Notemos que $x_{j,A} = -x_{j,A^-}$

Para $m < j$

$$\begin{aligned} x_j(m+2) &\leq x_j(m+1) \Longleftrightarrow (m+1)\lambda P_\lambda(U_{m+1}) \geq P_\lambda(U_m) \\ \Longleftrightarrow \quad (m+1)\lambda P_\lambda(U_m) + (m+1)P_\lambda(\{m+1\}) &\geq P_\lambda(U_m) \end{aligned}$$

Para $m \geq j$

$$\begin{aligned} & x_j(m+2) \leq x_j(m+1) \\ \Longleftrightarrow \quad & \frac{m+1}{\lambda} \sum_{k=m+2}^{\infty} \frac{e^\lambda \lambda^k}{k!} \leq \sum_{k=m+1}^{\infty} \frac{e^\lambda \lambda^k}{k!} \\ \Longleftrightarrow \quad & (m+1) \sum_{k=m+2}^{\infty} \frac{\lambda^{k-1}}{k!} \leq \sum_{k=m+1}^{\infty} \frac{\lambda^k}{k!} \end{aligned}$$

Lo cual se verifica al comparar término a término.

Por lo tanto, $x_j(n)$ es positiva y decreciente para $n \geq (j+1)$ y es negativa y decreciente para $n \leq j$. Así el único valor positivo tomado por $x_j(j+1) - x_j(m)$ es

$$\begin{aligned} x_j(j+1) - x_j(m) &= \frac{e^{-\lambda}}{\lambda}\left[\sum_{r=j+1}^{\infty} \frac{\lambda^r}{r!} + \sum_{r=1}^{j}\left(\frac{\lambda^r}{r!}\frac{r}{j}\right)\right] \\ &\leq \frac{e^{-\lambda}}{\lambda}\left[\sum_{r=1}^{\infty} \frac{\lambda^r}{r!}\right] \qquad \left(\text{pues } \frac{r}{j} \leq 1\right) \\ &= \frac{e^{-\lambda}}{\lambda}(e^\lambda - 1) \\ &= \lambda^{-1}(1 - e^{-\lambda}) \end{aligned}$$

De aquí que para cualquier $A \subset \mathbb{Z}^+$

$$\begin{aligned} x_{\lambda,A}(m+1) - x_{\lambda,A}(m) &= 1_{\{m\in A\}}\{x_m(m+1) - x_m(m)\} \\ &+ \sum_{j\in A, j\neq m} \{x_j(m+1) - x_j(m)\} \end{aligned}$$

donde sólo el primer sumando es positivo, por lo tanto

$$\begin{aligned}\sup_{m\in\mathcal{Z}^+}\max_{A\in\mathcal{Z}^+}\{x_{\lambda,A}(m+1)-x_{\lambda,A}(m)\} &= \sup_{m\in\mathcal{Z}^+}\{x_m(m+1)-x_m(m)\}\\ &\le \lambda^{-1}\left(1-e^{-\lambda}\right).\end{aligned}$$

Notemos que $\lambda^{-1}\left(1-e^{-\lambda}\right)\le\lambda^{-1}\lambda=1$, además $\lambda^{-1}\left(1-e^{-\lambda}\right)\le\lambda^{-1}$.

También que, como $x_{\lambda,A}=-x_{\lambda,A^-}$ entonces las diferencias tienen los mismos valores absolutos cuando tomamos A^-. Por lo tanto, basta considerar $A\subset\mathcal{Z}^+$. Así

$$\|\Delta x\|\le\min\{1,\lambda^{-1}\}.$$

Con lo cual concluimos la demostración del lema.

□

Lema 2

Con las mismas hipótesis de lema anterior, si $h(w)=1_{(w=0)}-e^{-\lambda}$, entonces

$$\|x\|=(1-e^{-\lambda})/\lambda.$$

Demostración:

El punto inicial de la demostración es la observación de que, si $E[h(Z)]=0$ entonces

$$\begin{aligned}(Sh)(w+1) &= -\lambda^{-1}P(Z=w)^{-1}E[h(Z);Z\le w]\\ &= -\lambda^{-1}P(Z=w)^{-1}\{E[h(Z)1_{(Z\le w)}]-E[h(Z)]E[1_{(Z\le w)}]\}\\ &= -\lambda^{-1}P(Z=w)^{-1}cov(h(Z),1_{(Z\le w)})\end{aligned}$$

Es fácil comprobar que para $k\ge 0$, si $h(w)=1_{(Z\le w)})-P(Z\le k)$ entonces

$$cov(h(Z),1_{(Z\le w)}=P(Z\le\min\{k,w\})-P(Z\le k)P(Z\le w).$$

Ahora:

$$\frac{d}{d\lambda}P(Z\le j) = \sum_{k=0}^{j}\frac{1}{k!}\frac{d}{d\lambda}\left(e^{-\lambda}\lambda^k\right)$$

$$\begin{aligned}
&= \sum_{k=0}^{j} \frac{1}{k!}\left(k\lambda^{k-1}e^{-\lambda} - \lambda^k e^{-\lambda}\right) \\
&= \sum_{k=1}^{j} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} - \sum_{k=0}^{j} \frac{e^{-\lambda}\lambda^{k}}{k!} \\
&= \frac{e^{-\lambda}\lambda^{j}}{j!} \;=\; -P(Z=j).
\end{aligned}$$

Lo cual nos lleva a la siguiente

Afirmación:

$$\begin{aligned}
P(Z \leq j) &= 1 - \int_0^{\lambda} e^{-v}\frac{v^j}{j!}dv \\
&= \int_{\lambda}^{\infty} e^{-v}\frac{v^j}{j!}dv
\end{aligned}$$

La primera igualdad se sigue de la observación 6 y la segunda del hecho de que

$$\int_0^{\infty} e^{-v}\frac{v^j}{j!}dv = 1.$$

Combinando todos estos ingredientes para el caso $k = 0$ (y entonces $w = 0$) tenemos que

$$\begin{aligned}
x(1) &= -\lambda^{-1}P(Z=0)[P(Z=0) - P(Z=0)^2] \\
&= -\lambda^{-1}[1 - P(Z=0)] \\
&= -\lambda^{-1}\int_0^{\lambda} e^{-v}dv \\
&= -\left(\frac{1-e^{-\lambda}}{\lambda}\right).
\end{aligned}$$

Así

$$\frac{1-e^{-\lambda}}{\lambda} = -x(1) > -x(2) > \cdots > 0.$$

Y por lo tanto

$$\|x\| = \frac{1-e^{-\lambda}}{\lambda}.$$

□

Apéndice 2.

El objetivo de este anexo es presentar los conceptos básicos de teoría de gráficas que son utilizados durante la exposición de algunos de los ejemplos de gráficas aleatorias que son tratados en este trabajo. Estos conceptos se pueden encontrar en [4] o en [10].

Una ***gráfica*** es una pareja de conjuntos X, A, donde X es un conjunto de puntos llamados *vertices o nodos* y A es un conjunto de la forma

$$A = \{\{x, y\} : x, y \in X\}.$$

El conjunto A puede ser pensado como el conjunto de lineas vértices en X. Usualmente denotamos a la gráfica como $G = [X, A]$.

Si los elementos de A son pares ordenados, es decir, las lineas tienen una dirección, éstos son llamados *arcos* y se dice que la gráfica G es *dirigida*. Si no tienen dirección se llaman *aristas* y G es *no dirigida*.

Un arco a puede representarse como la pareja $a = (x, y), x, y \in X$. x es llamado *vértice o extremo inicial* de a y y es el *vértice o extremo final* de a.

En una gráfica dirigida $G = [X, A]$, se llama *sucesor* del vértice $x \in X$ a todo vértice $y \in X$ tal que existe $(x, y) \in A$. Se llama *predecesor* de x a todo vértice $y \in X$ tal que existe $(y, x) \in A$.

También puede definirse tanto para gáficas dirigidas como no dirigidas los siguientes conceptos. Un vértice y es *vecino* de un vértice x si existe la arista $\{x, y\} \in A$. Si la gráfica es dirigida, $y \in X$ es vecino de $x \in X$ si y es sucesor o predecesor de x. Algunas veces se dice que x y y son *adyacentes*.

Una *cadena* es una secuencia de aristas (o arcos) $a_1, a_2, \ldots, a_q$ donde toda a_i está conectada a a_{i-1} por un extremo y a a_{i+1} por el otro. Si la cadena sólo cuenta con dos aristas éstas deben estar conectadas por uno de sus extremos. Un *ciclo* es una cadena $x_1, x_2, \ldots, x_q$ donde $x_1 = x_q$.

Se conoce como grados de un vértice a lo siguiente. Sea $G = [X, A]$ una gráfica dirigida y sea $x \in X$. El *grado exterior* de x es el número de arcos que tienen a x como vértice inicial. Se denota como $g^+(x)$. El *grado interior* de x es el número de arcos que tienen a x como vértice final. Se denota como $g^-(x)$.

El grado de $x \in X$ es el número de arcos que tienen a x como uno de sus extremos. Se denota como $g(x)$.

Existen ciertos subconjuntos de gráficas que son de utilidad y que definiremos a continuación. Sea $G = [X, A]$ una gráfica. Una *gráfica parcial* de G es la gráfica $G_p = [X, A_p]$, donde $A_p \subset A$, es decir, es una gráfica formada por todos los vértices originales y sólo algunas aristas o arcos. Una *subgráfica* de g es una gráfica $G_s = [X_s, A_s]$, donde $X_s \subset X, A_s \subset A$ y $(x, y) \in A_s$ si y sólo si $x, y \in X_s$ y $(x, y) \in A$, es decir, es una gráfica que consta de un subconjunto de los vértices de G y todas las aristas o arcos de G que unen a los vértices de tal subconjunto y que originalmente estaban en A.

Una gráfica $G = [X, A]$ es *conexa* si para todo par de vértices $x, y \in X$ existe una cadena que los une.

Finalmente daremos el concepto de gráfica aleaoria: Sea n un entero positivo y $0 \leq p \leq 1$. La *gráfica aleatoria* $G(n, p)$ es un espacio de probabilidad sobre el conjunto de gráficas con vértices $\{1, 2, \ldots, n\}$ determinado por $P(\{x, y\} \in G) = p$, con esos eventos mutuamente independientes.

Bibliografía

[1] Arratia, R., Goldstein, L. & Gordon, L. *Poisson approximations and the Chen-Stein method.* Statistical Science, 1990, Vol. 5, No. 4, 403 - 434.

[2] Arratia, R., Goldstein, L. & Gordon, L. *Two moments suffice for Poisson approximations: The Chen-Stein method.* The annals of Probability. 1989, Vol. 17, No. 1, 9 - 25.

[3] Barbour, A. D., & Eagleson, G. K. *Poisson aproximation for some statistics based on exchangeable trials.* Adv. App. Prob. 1983, **15**, 585 - 600.

[4] Bazaraa, M. S. and Jarvis,G. J. *Linear Programming and Network Flows.* Jhon Wiley & sons. New York. 1977

[5] Bergamini, David. *Matemáticas.* Colección científica de TIME-LIFE. Ediciones culturales Internacionales. 1988.

[6] Chen, L. H. Y. *Poisson approximation for dependent trials.* The annals of Probability. 1975, **3**, 534 - 545.

[7] Durret, R. *Probability: Theory and examples.* Wadsworth & Brooks/Cole. Statistics/Probability Series. 1991.

[8] Feller, W. *An introduction to Probability theory and its applications.* Vol. 1, 3ª edic. Jhon Wiley & sons. New York. 1968.

[9] Feller, W. *An introduction to Probability theory and its applications.* Vol. 2, 2ª edic. Jhon Wiley & sons. New York. 1971.

[10] Hillier, F. S. & Lieberman, G. J. *Introduction to Operations Research.* Holden-Day, Inc. 1980.

[11] Karlin, S. & Taylor, H. M. *A first course in Stochastic Processes.* Academic Press, Inc., 1984.

[12] Ross, Sheldon M. *A first course in Probability.* Maxwell Macmillan International Editions. 1989.

[13] Ross, Sheldon M. *Stochastic Processes.* John Wiley & sons, 1993.

[14] Stein, C. M. *A bound for the error in the Normal approximation to the distribution of a sum of dependent random variables.* Proc. Sixth Berkeley Symp. Math. Statist. Probab. 1971, **3**, 583 - 602.

[15] Takács, L. *On the limit distribution of the number of the cycles in a random graph.* J. Appl. Prob. 1988 **26**, 359 - 376.

[16] Trujillo Cortez, R. *Series aleatorias.* Tesis. 1984. Facultad de Ciencias. UNAM.