UNIVERSIDAD NACIONAL AUTONOMA DE MÉXICO

FACULTAD DE CIENCIAS

CONSIDERACIONES SOBRE LA FUNCIÓN TASA Y ALGUNAS APLICACIONES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A:

ABEL CAMACHO GUARDIAN

DIRECTOR DE TESIS:
DRA. ANA MEDA GUARDIOLA

2009

UNAM Universidad Nacional Autónoma de México

Biblioteca Central

Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno.
Camacho
Guardian
Abel
55541651
Universidad Nacional Autonoma de Mexico
Facultad de Ciencias
Matemáticas
303730800

2. Datos del asesor
Dra
Meda
Guardiola
Ana

3. Datos del sinodal 1
Dra.
María Asunción Begoña
Fernández
Fernández

4. Datos del sinodal 2
Dr.
Gerónimo Francisco
Uribe
Bravo

5. Datos del sinodal 3
M. en C.
Juan Martín
Barrios
Vargas

6. Datos del sinodal 4
M. en C.
Gerardo
Rubio
Hernández

7. Datos de la tesis
Consideraciones sobre la función tasa y algunas aplicaciones
55p
2009

Agradecimientos

Primero que nada, este trabajo no hubiera sido posible sin el apoyo de mi familia, es decir a mis padres (Beatriz, Abel), mi abuelo (Abel) y a mi hermano (Arturo). La carrera se la debo a ellos.

Les estoy profundamente agradecido a Guisselle y Yazmin por la amistad.

Esta tesis no sería la misma sin Yazmin, quien estuvo la mayoría del tiempo de la realización de esta en la misma mesa de la Biblioteca, ni Guisselle quien estuvo del otro lado del teléfono.

A mi sinodal y maestra Ana Meda, quien le debo mucho en mi formación como matemático

A mis sinodales Begoña, Gerónimo, Juan Barrios y Gerardo que me hicieron observaciones valiosas para mejorar mí trabajo.
Un agradecimiento al sinodal Juan por ayudarme en todo lo referente de computación.

Índice general

1. Introducción **3**
1.1. Introducción 3

2. Preliminares **5**
2.1. Preliminares 5
2.1.1. Cambios de Medida 6
2.1.2. Desigualdades 9
2.1.3. Leyes de los grandes números 12

3. Teorema Principal **15**

4. Observaciones de la Función Tasa **21**
4.1. Función Tasa 21
4.1.1. Ejemplos 34
4.1.2. Teorema 41
4.1.3. Aplicación 44

5. Grandes Desviaciones para Proceso de Ramificación **49**
5.1. Galton- Watson 52
5.2. Procesos de ramificación y grandes desviaciones 54

Capítulo 1

Introducción

1.1. Introducción

El presente trabajo trata de las grandes desviaciones. Más precisamente, el primer interés en nuestro trabajo es comprender mejor el teorema débil de los grandes números en el sentido siguiente. Por el teorema débil de los grandes números podemos decir, si tenemos variables aleatorias $\{X_i\}_{i=1}^{\infty}$ independientes e idénticamente distribuidas con función generadora de momentos, $M(\Theta)$, finita en algún intervalo abierto del cero y $a > E(X)$, entonces

$$P(\frac{X_1 + ... + X_n}{n} \geq a) \ converge\ a\ cero,\ conforme\ n\ tiende\ a\ \infty \tag{1.1}$$

sin embargo no nos dice nada sobre el comportamiento asintótico de esta sucesión.

En el capítulo dos veremos que la sucesión converge de forma exponencial: $a_n := P(\frac{X_1+...+X_n}{n} \geq a) = e^{-nl(a)+o(n)}$, donde $l(a) = sup_\Theta a\Theta - \ln(M(\Theta))$. En el teorema principal de ese capítulo se supone que el supremo es un máximo.

El capítulo tres busca entender mejor el significado de que el supremo sea máximo y conocer propiedades de la función tasa: $l(\cdot) : (E(X), \infty) \longrightarrow \Re$; esto permitirá extender el teorema del capítulo dos sobre los conjuntos y eliminar la restricción del máximo. También veremos cómo la existencia del máximo esta íntimamente relacionada con la continuidad de la función $l(\cdot)$.

El capítulo cuatro relaciona el trabajo de los capítulos anteriores con el proceso de ramificación, aquí generalizaremos el problema de extinción planteado en el modelo de Galton-Watson.

Como se verá a lo largo del trabajo estudiar a la función tasa en vez de el teorema de Cramér como usualmente vienen en los libros, permite:

(I) Pasar de problemas sobre sucesiones de probabilidades a problemas sobre continuidad de la función tasa.

(II) Al adquirir ejemplos de funciones tasa para distintas variables aleatorias, tenemos ejemplos sobre el distinto comportamiento asintótico de (1.1), es decir sobre el teorema de Cramér.

(III) Las propiedades de la función tasa están ligadas con el supremo de la función: $a\Theta - \ln(M(\Theta))$ que es una función cóncava; por lo tanto es sencillo decir cuándo existe un máximo.

(IV) Sólo requerimos de un curso de teoría de la medida y conocimientos de cálculo.

Capítulo 2

Preliminares

2.1. Preliminares

En esta sección presentamos el material básico para entender el trabajo posterior.

Muchas de las proposiciones que aquí se requieren tienen el mismo principio básico de demostración que a continuación enunciaremos y que llamaremos como "Principio estándar"

Principio estándar: Si queremos demostrar una proposición para una clase de funciones, hacemos:

i) Mostramos que la proposición es válida para funciones indicadoras $f = I_A$ con $A \in \mathcal{F}$, y $\mathcal{F}$ es la σ-álgebra.

ii) Usamos la linealidad para mostrar que es válido para funciones simples.

iii) Con el teorema de la convergencia monótona mostramos que es válido para cualquier función positiva y medible.

iv) Finalmente usando que toda función f se puede ver como $f = f^+ - f^-$, con $f^+, f^- \geq 0$ vemos que la proposición es válida para toda función integrable.

Definición A. 1 *Sea X una variable aleatoria y $D = \{s \in \Re : E(e^{(sX)}) < \infty\}$. La función $M : D \longrightarrow \Re$, con $M(s) = E(e^{(sX)}) = \int e^{(sX)} dP$ es la función generadora de momentos.*

Teorema A. 1 *Sea X una variable aleatoria con función generadora de momentos $M : D \longrightarrow \Re$ y $(-b, b) \subset D$ para alguna $b > 0$; entonces se satisface lo siguiente: existen todos los momentos, $E(X^n)$ es finito y*

$$M(s) = \sum_{n=0}^{\infty} \frac{s^n}{n!} E(X^n) \quad con \quad s \in (-b, b) \tag{2.1}$$

en particular se cumplen:

$$M^n(0) = E(X^n).$$

Demostración.

Basta mostrar que en el intervalo $(-b, b)$ podemos sacar la serie $e^{(Xs)} = \sum_{n=0}^{\infty} \frac{(sX)^n}{n!}$ de la integral, es decir, podemos intercambiar esperanza por serie. Por hipótesis, e^{Xs} es P-integrable para todo $s \in (-b, b)$, por lo que también es válido para el valor absoluto, ya que $e^{|sX|} \leq e^{sX} + e^{-sX}$; en otras palabras, es válido para $\sum_{n=0}^{\infty} \frac{|sX|^n}{n!} \geq 0$. Haciendo usó del Teorema de Convergencia Dominada sobre la siguiente sucesión de funciones,

$$f_N(X) := \sum_{n=0}^{N} \frac{(sX)^n}{n!} \quad y \quad g(X) := e^{|sX|} \tag{2.2}$$

$$M(s) = E(e^{sX}) = E(Lim_{N \longrightarrow \infty} f_N) =$$

$$Lim_{N \longrightarrow \infty} E(f_N) = \sum_{n=0}^{\infty} \frac{s^n}{n!} E(X^n) \quad s \in (-b, b)$$

que nos permite concluir (2.1). De la integrabibilidad de $f_N(X)$ tenemos que $E(X^n)$ es finito para toda $n \in \mathbb{N}$. Como cualquier serie de potencias $M(s)$ tiene las derivadas de cualquier orden dentro de su radio de convergencia y se obtienen derivando término a término.

◇

2.1.1. Cambios de Medida

En esta subsección no requerimos que P sea una probabilidad, por lo que no lo supondremos. Usamos la notación P por qué más adelante solo trabajamos con espacios de probabilidad.

Definición A. 2 *Supongamos que g es una función medible, no negativa y definamos la medida P_g por:*

$$P_g(A) = \int_A g dP. \tag{2.3}$$

Decimos que P_g definida por 2.3 es una medida con densidad g relativo a la medida P

Teorema A. 2 *Sea P_g la medida con densidad g (rel a P), entonces*

$$\int f dP_g = \int f g dP \tag{2.4}$$

se satisface para toda función f no negativa. Más aún, f es integrable, no necesariamente no negativa, con respecto a P_g si y sólamente si fg es integrable con respecto a P, en cuyo caso se satisface la igualdad. Para f no negativa tenemos además lo siguiente: $\forall A \in \mathcal{F}$

$$\int_A f dP_g = \int_A fgdP. \tag{2.5}$$

Demostración.

Si la función es de la forma $f = 1_A$, entonces $\int f dP_g = \int_A 1 dP_g = \int fgdP$, de tal manera que la ecuación (2.4) se satisface. Si se da el caso que f es una función simple, (2.4) se sigue por linealidad. Para f no negativa y medible, tenemos que existe una sucesión creciente de funciones simples $\{f_n\}_{n=1}^{\infty}$ que converge a f y por lo anterior $\int f_n dP_g = \int f_n gdP$ $\forall n \in \mathbb{N}$ y por el teorema de convergencia monótona se tiene: $\int f dP_g = \int fgdP$. Obsérvese cómo ambas tienen que ser un real o infinito a la vez. Para el caso general usamos que una función f es integrabe si y sólo si $|f|$ es integrable, vemos que f es P integrable si y sólo si f es P_g integrable y que $f = f^+ - f^-$. Para el resultado (2.5) remplazamos f por $f1_A$.

$\diamond$

Nótese cómo la demostración sigue el principio estándar y nunca se usó que $P(\Omega) = 1$.

Definición A. 3 CAMBIO DE VARIABLE.

Sean $(\Omega, \mathcal{F})$ y $(\Omega', \mathcal{F}')$ dos espacios medibles y sea $T : \Omega \longrightarrow \Omega'$ una función medible. Para una medida P sobre $\mathcal{F}$, podemos definir una medida $P^{T^{-1}}$ sobre $\mathcal{F}'$ de la siguiente manera:

$$P^{T^{-1}}(A') = P(T^{-1}(A)) \qquad \forall A' \in \mathcal{F}'.$$

Teorema A. 3 *Sea f una función no negativa, entonces tenemos*

$$\int_\Omega f(T) dP = \int_{\Omega'} f dP^{T^{-1}}. \tag{2.6}$$

Con f no necesariamente es no negativa, es integrable con respecto a $P^{T^{-1}}$ si y sólamente si $f(T)$ es integrable con respecto a P, en cuyo caso se satisface la igualdad de arriba. Para funciones no negativas adicionalmente:

$$\int_{T^{-1}(A)} f(T) dP = \int_A f dP^{T^{-1}}. \tag{2.7}$$

Demostración.

De nuevo se seguirá el principio estándar. Si $f = 1_{A'}$, entonces $f(T) = 1_{T^{-1}}(A')$ y se sigue (2.6). Para funciones simples se sigue por linealidad de la integral.Y del mismo modo que en la demostración anterior, para $f \geq 0$ y medible tenemos una sucesión creciente $\{f_n\}_{n=1}^{\infty}$ de funciones simples que convergen a f, entonces $\{f_n(T)\}_{n=1}^{\infty}$ converge a $f(T)$ y por el Teorema de convergencia Monótona se vale (2.6). Aplicando $|f|$ tenemos que f es $P_{T^{-1}}$ integrable si y sólamente si $f(T)$ es P, integrable, ya que la integrabilidad de ambas se da cuando para $|f|$ su integral no es infinito. La igualdad de (2.6) se sigue de descomponer a $f = f^+ - f^-$. Por último, para (2.7) remplazamos f por $f1_{A'}$.

$\diamond$

Aquí tampoco se usó que $P(\Omega) = 1$. Nótese como tenemos un cambio de variable cuando $(\Omega', \mathcal{F}') = (\Omega, \mathcal{F})$.

A continuación seremos menos estrictos en enunciar los resultados más generales. La intención es mostrar cierta relación entre las dos definiciones de esta sección, esperando que de este modo queden más claras y el entendimiento de una contribuya a la comprensión de la otra. Sea f una función λ-integrable y mayor que cero, además que satisfaga que $F(x) := \int_{-\infty}^{x} f dx$, estamos pidiendo que sea Riemann integrable. F es continua y estrictamente creciente.

Observación 1, $F^{-1} : \Re \longrightarrow \Re$ es continua y estrictamente creciente.

Observación 2, $\lambda_f((a,b]) = \int_{(a,b]} 1 f d\lambda = \int_{a^+}^{b} f(y)dy = F(b) - F(a)$, esta última igualdad por continuidad de F.

Observación 3, $\lambda^F((a,b]) = \int_{(a,b]} 1 d\lambda^F = \int_{(F(a),F(b)]} 1 d\lambda = F(b) - F(a)$.

Observación 4 Las integrales de las observaciones 1 y 2 pueden sustituirse por los intervalos (a, ∞) y $(-\infty, b]$ esto $\forall\, a, b \in \Re$.

Observación 5, por las observaciones 2 y 3 y que $F(F^{-1})(y) = y$ se sigue que $(\lambda_f)^{F^{-1}} = \lambda$ y $(\lambda^{F^{-1}})_f = \lambda$, lo cual nos permite escribir $\lambda_f^{F^{-1}} = \lambda$. De lo anterior tenemos una relación bastante bonita.

Sea g una función λ-integrable y Riemann integrable en $(a, b]$, entonces,

$$\begin{gathered}
\int_{(a,b]} g(F) f d\lambda = \int g(F) f 1_{(a,b)} d\lambda = \int g(F) 1_{(a,b)} d\lambda_f = \\
\int g(F) 1_{(a,b)} (F^{-1}(F)) d(\lambda_f) = \int g 1_{(a,b)} (F^{-1})(F) d(\lambda_f) = \\
= \int g 1_{(a,b)} (F^{-1}) d\lambda_f^{F^{-1}} = \int g 1_{(a,b)} (F^{-1}) d\lambda = \\
\int_{(F(a),F(b)]} g d\lambda
\end{gathered}$$

Y por ser g y $g(F)f$ Riemann integrables,

$$\int_a^b g(F(x)) F' dx = \int_{F(a)}^{F(b)} g(y) dy \tag{2.8}$$

Lo que corresponde a nuestro famoso cambio de variable de cálculo.

2.1.2. Desigualdades

Proposición A. 1 DESIGUALDAD DE TCHEBYSHEV.

Sea (Ω, F, P) un espacio de probabilidad y $\phi : [0, \infty] \longrightarrow [0, \infty]$ una función no decreciente y mayor que cero en $(0, \infty)$ dada. Sea X una variable aleatoria tal que $\phi(X)$ tiene esperanza, entonces para toda $a > 0$ tenemos:

$$P(|X| \geq \alpha) \leq \frac{E(\phi(|X|))}{\phi(\alpha)}.$$

Demostración.

Ocurre $P(|X| \geq \alpha) \leq \frac{E(\phi(|X|))}{\phi(\alpha)}$ si y sólamente si $\phi(\alpha)P(|X| \geq \alpha) \leq E(\phi(|X|))$. Y el lado izquierdo es igual a $E(\phi(\alpha)1_{\{|X| \geq \alpha\}})$ que por monotonía de la integral es menor o igual que $E(\phi(|X|)1_{\{|X| \geq \alpha\}})$ y de nuevo por monotonía: $E(\phi(|X|))$.

$\diamond$

Algunas consecuencias son las siguientes:

$P(|X| \geq \alpha) \leq \frac{E(|X|)}{\alpha}$ (Desigualdad de Markov)

Si X tiene momento de orden n, $P(|X| \geq \alpha) \leq \frac{E(|X|^n)}{\alpha^n}$.

Con $Y = (X - E(X))$ tenemos $P(|X - E(X)| \geq \alpha) \leq \frac{Var(X)}{\alpha}$.

Proposición A. 2 DESIGUALDAD DE JENSEN. *Sea $\phi : \Re \longrightarrow \Re$ una función convexa y sea X, $\phi(X)$ variables aleatorias con esperanza finita, entonces*

$$\phi(E(X)) \leq E(\phi(X)).$$

Antes cubriremos un prerequisito sobre funciones convexas.

Teorema A. 4 *Sea $\phi : I \longrightarrow \Re$ una función convexa, entonces*

i)Si $x \in I^o$, existen las derivadas derecha e izquierda y satisfacen:

$$D^-(\phi)(x) \leq D^+(\phi)(x).$$

Demostración.

Sea $x \in I^o$, entonces existen $\lambda_1^{'}, \lambda_2^{'} > 0$ tales que si $\lambda_i < \lambda_i^{'}$ con $i = 1, 2$, $x - \lambda_2, x + \lambda_1 \in I^o$, además

$$x = \frac{\lambda_1}{\lambda_1+\lambda_2}(x - \lambda_2) + \frac{\lambda_2}{\lambda_1+\lambda_2}(x + \lambda_1).$$

Pero la convexidad de ϕ nos lleva:

$$\phi(x) \leq \frac{\lambda_1}{\lambda_1+\lambda_2}\phi(x - \lambda_2) + \phi(x) \leq \frac{\lambda_2}{\lambda_1+\lambda_2}\phi(x + \lambda_1)$$

multiplicando por $\lambda_1 + \lambda_2$, tenemos

$$(\lambda_1 + \lambda_2)\phi(x) \leq \lambda_1\phi(x - \lambda_2) + \phi(x) \leq \lambda_2\phi(x + \lambda_1).$$

Que nos da:

$$\lambda_1\phi(x) - \lambda_1\phi(x - \lambda_2) \leq \lambda_2\phi(x + \lambda_1) - \lambda_2\phi(x).$$

Dividiendo entre $\lambda_1\lambda_2$

$$\tfrac{\phi(x)-\phi(x-\lambda_2)}{\lambda_2} \leq \tfrac{\phi(x+\lambda_1)-\phi(x)}{\lambda_1}.$$

En resumen,

Si $\lambda_1, \lambda_2 > 0$ con $x - \lambda_2, x + \lambda_1 \in I^o$, se satisface:

$$\frac{\phi(x) - \phi(x - \lambda_2)}{\lambda_2} \leq \frac{\phi(x) - \phi(x + \lambda_1)}{-\lambda_1}. \tag{2.9}$$

Ahora, si $0 < \lambda_1 < \lambda_2$ con $x + \lambda_2, x + \lambda_1 \in I^o$, tenemos lo siguiente:

$$\phi(x + \lambda_1) = (\tfrac{\lambda_1}{\lambda_2}(x + \lambda_2) + (1 - \tfrac{\lambda_1}{\lambda_2})x)$$

que por convexidad:

$$\phi(x + \lambda_1) \leq \tfrac{\lambda_1}{\lambda_2}\phi(x + \lambda_2) + (1 - \tfrac{\lambda_1}{\lambda_2})\phi(x).$$

Que a su vez se reduce:

$$\phi(x + \lambda_1) - \phi(x) \leq \tfrac{\lambda_1}{\lambda_2}(\phi(x + \lambda_2) - \phi(x)).$$

Dividiendo entre λ_1:

$$\frac{\phi(x) - \phi(x + \lambda_1)}{\lambda_1} \leq \frac{\phi(x) - \phi(x + \lambda_2)}{\lambda_2}. \tag{2.10}$$

De forma análoga tenemos:

Si $0 < \lambda_1 < \lambda_2$ con $x + \lambda_2, x + \lambda_1 \in I^o$, entonces

$$\frac{\phi(x) - \phi(x + \lambda_2)}{-\lambda_2} \leq \frac{\phi(x) - \phi(x + \lambda_1)}{-\lambda_1} \tag{2.11}$$

De las ecuaciones 2.9, 2.10 y 2.11 concluimos lo siguiente.

$$\exists Lim_{\lambda \longrightarrow 0^+}\frac{\phi(x) - \phi(x + \lambda)}{-\lambda} = inf_{\lambda > 0}\frac{\phi(x) - \phi(x + \lambda)}{-\lambda} = D^+(\phi)(x) \tag{2.12}$$

$$\exists Lim_{\lambda \longrightarrow 0^+} \frac{\phi(x) - \phi(x-\lambda)}{\lambda} = sup_{\lambda>0} \frac{\phi(x) - \phi(x-\lambda)}{\lambda} = D^-(\phi)(x) \quad (2.13)$$

y finalmente:

$$D^-(\phi)(x) \leq D^+(\phi)(x). \quad (2.14)$$

$\diamond$

Observaciones.

Ambas derivadas parciales son reales.

Una función convexa es continua en el interior de su dominio.

La siguiente observación nos dice que para todo punto en el interior de I, existe una recta que está por abajo de la función ϕ en todo I.

Observación A. 1 *Sea c tal que $c \in [D^-(\phi)(x), D^+(\phi)(x)]$ y $b := \phi(x) - cx$, definamos a la recta $l(y) = cy + b$, entonces $l(y) \leq \phi(y)$ para todo $y \in I$.*

Demostración.

Sea $y > x$, es decir $y = x + \lambda$ para alguna $\lambda > 0$ $l(y) \leq \phi(y)$ si y sólamente si $cy + (\phi(x) - cx) \leq \phi(y)$ si y sólamente si $cy - cx \leq \phi(y) - \phi(x)$ que por definición de y tenemos $c(x+\lambda) - cx \leq \phi(x+\lambda) - \phi(x)$ dividiendo entre λ $c \leq \frac{\phi(x+\lambda)-\phi(x)}{\lambda} = \frac{\phi(x)-\phi(x+\lambda)}{-\lambda}$ Pero por como tomamos a c y por propiedades de $D^+(\phi)(x)$ se sigue que:

$$c \leq D^+(\phi)(x) \leq inf_{\lambda>0} \tfrac{\phi(x)-\phi(x+\lambda)}{-\lambda}.$$

Análogamente:

Sea $y < x$, es decir $y = x - \lambda$ para alguna $\lambda > 0$, entonces $l(y) = cy + b \leq \phi(y)$ $\diamond$

Del Teorema A.4 y de la observación A.1, se sigue la Desigualdad de Jensen.
Demostración.(Desigualdad de Jensen)
Definamos a $x = E(X)$, por hipótesis $x \in \Re$, luego entonces sea $c := D^+(\phi)(x)$ y $b := \phi(x) - cx$ de la observación anterior se sigue: $cy + b \leq \phi(y)$ y con $y = x$ la desigualdad es igualdad.

$$\phi(E(X)) = c(E(X)) + b = c(E(X)) + b(E(1)) = E(cX + b).$$

Y por monotonía de la esperanza,

$$E(cX + b) \leq E(\phi(X)).$$

Que concluye la desigualdad de Jensen. $\diamond$

Como una función cóncava φ es menos una función convexa, $-\varphi$, tenemos inmediatamente que $E(\varphi(X)) \leq \varphi(E(X))$.

Como ejemplos, $e^{E(X)} \leq E(e^X)$, $(E(X))^2 \leq E(X^2)$, $E(\ln(|X|)) \leq \ln(E(|X|))$.

2.1.3. Leyes de los grandes números

Teorema A. 5 BOREL-CANTELLI, ERDÖS-RÉNYI

Sea $(A_n)_{n\geq 1}$ *una sucesión de eventos en el espacio de probabilidad* (Ω, F, P), *entonces*

$$\sum_{n\geq 1} P(A_n) < \infty \Rightarrow P(Lim\, supA_n) = 0. \tag{2.15}$$

Si los eventos son independientes, entonces también es válido

$$\sum_{n\geq 1} P(A_n) = \infty \Rightarrow P(Lim\, supA_n) = 1.$$

Debe entenderse a $Lim\, supA_n := \cap_{n=1}^{\infty} \cup_{k\geq} A_k$.

Del teorema sólo usamos (2.15) en la aplicación sobre la ley fuerte de los grandes números y en el teorema de Biggins, por lo que sólo nos concentraremos en la demostración de esta parte.

Demostración.

Como $\cup_{m\geq n} A_m \downarrow \cap_{n\in\mathbb{N}} \cup_{m\geq n} A_m$ y$P(A_1) \leq 1 < \infty$ se sigue,

$$P(Lim\, supA_n) = Lim_{n\longrightarrow\infty} P(\cup_{m\geq n} A_n) \leq Lim_{n\longrightarrow\infty} \sum_{m=n}^{\infty} P(A_m) = 0$$

$\diamond$

Definición A. 4 CONVERGENCIA CASI-SEGURAMENTE

La sucesión $\{X_n\}_{n=1}^{\infty}$ *de variables aleatorias converge P-casi seguramente (P-c.s) a la variable aleatoria* x, *cuando*

$$P(w \in \Omega : \; Lim_{n\longrightarrow\infty} X_n(w) \;\; existe \;\; y \;\; es \;\; igual \;\; a \;\; x(w)) = 1.$$

Definición A. 5 CONVERGENCIA EN PROBABILIDAD *Una sucesión de variables aleatorias* $\{X_n\}$ *decimos que converge en probabilidad a la variable aleatoria* X,*si* $\forall\, \varepsilon > 0$ *tenemos;*

$$Lim_{n\longrightarrow\infty} P(|X_n - X| \geq \varepsilon) = 0.$$

Definición A. 6 LEY DE LOS GRANDES NÚMEROS

Sea $\{X_n\}$ *una serie de variables aleatorias integrables. Diremos que para la sucesión* $\{X_n\}$ *es válida la ley débil de los grandes números, cuando*

$Y_n := \frac{1}{n} \sum_{i=1}^{n} (X_i - E(X_i))$

$$Y_n \longrightarrow 0 \quad \text{en probabilidad}$$

y la sucesión $\{X_n\}$ satisface la ley fuerte de los grandes números, cuando

$$Y_n \text{ converge a } 0 \ P\text{-c.s}$$

Teorema A. 6 *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, si*

$$Lim_{n\longrightarrow\infty}\frac{1}{n^2}\sum_{i=1}^{n}Var(X_i) = 0 \tag{2.16}$$

entonces para la sucesión $\{X_n\}$ es válida la ley débil de los grandes números.

Demostración.

De la ecuación (2.16) tenemos que, en particular, para toda X_i $i \in \mathbb{N}$ la varianza es finita. Por la desigualdad de Tchebyshev y por la Fórmula de Bienaymé (para variables aleatorias sin correlación dos a dos tenemos: $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i))$ para toda $\varepsilon > 0$:

$$P(|\frac{1}{n}\sum_{i=1}^{n}(X_i - E(X_i))| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}Var[\frac{1}{n}\sum_{i=1}^{n}(X_i - E(X_i))] =$$
$$\frac{1}{(n\varepsilon)^2}\sum_{i=1}^{n}Var(X_i) \longrightarrow 0$$

$\diamond$

Una observación es que la independencia pudo sustituirse por variables aleatorias sin correlación dos a dos.

Otra forma alternativa del teorema anterior es el siguiente corolario.

Corolario A. 1 *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes. Entonces es válida la ley débil de los grandes números para $\{X_n\}_{n\in\mathbb{N}}$ cuando se satisface alguna de las dos condiciones:*

(i) Las varianzas están acotadas por una constante c, $Var(X_n) \leq c < \infty$ para toda $n \in \mathbb{N}$.

(ii) El cuarto momento centrado está acotada por una c, es decir, $E((X_n - E(X_n)^4) \leq c < \infty \ \forall n \in \mathbb{N}$

Demostración.

Para (i) tenemos que,

$$\frac{1}{n^2}\sum_{i=1}^{n}Var(X_i) \leq \frac{1}{n^2}nc \longrightarrow 0$$

y se sigue del teorema anterior.

(ii) es un caso especial del inciso (i) por la siguiente razón,

$$Var(X_n) = E([X_n - E(X_n)]^2 1^2)) \leq E(X_n - E(X_n)]^4)^{\frac{1}{2}} \leq c^{\frac{1}{2}}.$$

Lo anterior es válido por la desigualdad de Hölder con $p = q = 2$, que se reduce al inciso (i) con $c^{\frac{1}{2}}$.

$\diamond$

Ejemplo A. 1 *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes con $Var(X_n) \leq c$ y $E(X_n) = m$, entonces*

$$\frac{X_1 + ... + X_n}{n} \longrightarrow m \quad \textit{en Probabilidad.}$$

Una versión de la ley fuerte de los grandes números para variables aleatorias con función generadora de momentos definida en un interval del cero se demuestra en el capítulo 3, en la sección llamada Aplicación. Aquí sólo presentamos una versión más general.

Teorema A. 7 LEY FUERTE DE LOS GRANDES NÚMEROS PARA VARIABLES INDEPENDIENTES E IDÉNTICAMENTE DISTRIBUIDAS

Sea $\{x_n\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $m := E(x) < \infty$, entonces se cumple la ley fuerte de los grandes números para la sucesión $\{X_n\}$, es decir

$$\tfrac{1}{n} \textstyle\sum_{i=1}^{n} x_i \longrightarrow m \ \textit{P-c.s.}$$

Capítulo 3

Teorema Principal

El teorema de este capítulo trata a los eventos donde $X_1 + ... + X_n$ difieren de $nE(X)$ por algo de orden n; en este capítulo solo tratamos eventos de la forma $A_n = \{X_1 + ...X_n \geq an\}$ con $a > E(X)$ y $P(X > a) > 0$ que por la ley débil de los grande números sabemos que $P(A_n)$ converge a cero conforme n tiende a infinito. Aquí buscamos conocer de qué manera se da esta convergencia a cero. Veremos cómo este decaimiento es exponencial con n.

$$Lim_{n \longrightarrow \infty} \frac{1}{n} \ln \Big(P(X_1 + ... + X_n \geq an) \Big) = -l(a)$$

Teorema 1 *Consideremos una sucesión $\{X_i\}_{i=1}^{\infty}$ de variables aleatorias independientes e idénticamente distribuidas. Para cualquier $a > E(X_1)$ y n un entero positivo, tenemos:*

$$P(X_1 + ... + X_n \geq na) \leq e^{-nl(a)} \tag{3.1}$$

Además, suponiendo que $M(\Theta) < \infty$ en alguna vecindad del 0 y que $l(a) = -\ln(E(e^{\Theta^*(X-a)}))$ se alcance en algún Θ^* dentro de la vecindad, entonces, para cualquier $\varepsilon > 0$ existe un entero $m := m_\epsilon$ tal que si $n > m$,

$$P(X_1 + ... + X_n \geq na) \geq e^{-n(l(a)+\varepsilon)} \tag{3.2}$$

Las ecuaciones (1) y (2) implican

$$P(X_1 + ... + X_n \geq na) = e^{-n(l(a)+o(n))} \tag{3.3}$$

que también se puede ver como:

$$Lim_{n \longrightarrow \infty} \frac{1}{n} \ln \Big(P(\frac{X_1 + ... + X_n}{n} \in [a, \infty)) \Big) = -l(a) \tag{3.4}$$

Unos comentarios antes de la demostración.

Con el siguiente capítulo podremos decir lo siguiente:

I. Que exista Θ^* es equivalente a pedir que $P(X > a) > 0$, Observación 4.

II. $\Theta^* > 0$, Observación 1.

III. La existencia de Θ^* garantiza que hay una vecindad abierta en la que $l(\cdot)$ es continua y diferenciable, Observación 7.

IV. El caso $P(X > a) = 0$, donde no existe Θ^*, se divide a su vez en dos (a) $P(X = a) = 0$ y (b) $P(X = a) > 0$. El caso (a) tendremos que $l(a) = \infty$, por lo que las desigualdades se satisfacen trivialmente, todo se reduce a ceros. El caso (b) es más complicado y su demostración se encuentra en la proposición 10, su demostración requiere de este teorema pero además necesitamos conocer donde la función $l(\cdot)$ es continua. Un ejemplo de estos dos casos está en el ejemplo 2 (Binomial).

V. Por comodidad manejamos $a > E(X)$, pero como se puede apreciar, el caso $P(X \leq na)$ con $a < E(X)$ es simplemente $b := -a$, $b > E(-X)$, $P(-X \geq nb)$ y nuestra variable aleatoria es $-X$.

VI. La ecuación (3.4) se extenderá a otros conjuntos, Teorema 6.

Las siguientes proposiciones son demostradas en el siguiente capítulo, pero se requieren en esta demostración

(1) Sea X una variable aleatoria tal que su función generadora de momentos $M(\Theta) : D \longrightarrow \Re$ satisface $0 \in D^o$, $E(X) = 0$; entonces tenemos que $f(\Theta) \leq 0$ si $\Theta \leq 1$ y se satisface con igualdad cuando $\Theta = 0$. Proposición 3.

(2) Si $P(X > a) > 0$, entonces $l(a) = a\Theta^* + \ln(M(\Theta^*))$ y $\Theta^* \in U$ con U un intervalo abierto donde la función generadora de momentos es finita

Demostración.

Por la proposición 3 del siguiente capítulo basta probar el resultado cuando $E(X) = 0$. La proposición 3 nos permite decir,

$$inf_{\{\Theta \in \Re\}} e^{-\Theta a} M(\Theta) = inf_{\{\Theta \geq 0\}} e^{-\Theta a} M(\Theta)$$

Cota superior. Sea $\Theta \geq 0$

$P(X_1 + ... + X_n \geq na) \leq P(\Theta(X_1 + ... + X_n) \geq \Theta na) =$

$= P(e^{\Theta(X_1+...+X_n)} \geq e^{\Theta na}) \leq e^{-\Theta na} E(e^{\Theta(X_1+...+X_n)}) =$

$= (e^{-\Theta a} E(e^{\Theta X}))^n$ esto para toda $\Theta \geq 0$

La primera desigualdad es igualdad si $\Theta > 0$; la primera igualdad es por que $e^{(\cdot)}$ es una función estrictamente creciente, la segunda desigualdad es por Tchebychev y la última igualdad es por independencia.
Tomando el ínfimo y por continuidad de la función $(\cdot)^n$,
$P(X_1 + ... + X_n \geq na) \leq (e^{-l(a)})^n$

Dado la relación 1-1, entre las funciones continuas por la derecha y no decrecientes, F con $F(-\infty) = 0$, $F(\infty) = 1$, y las medidas de probabilidad, P, en los subconjuntos Borelianos de la recta real no seremos muy estrictos en distinguir estas dos nociones.

En la cota inferior usaremos varios resultados del apéndice. El primero de ellos es el cambio de variables dado por el Teorema $A.2$, sea F la distribución de X. Definimos a G como sigue:

$$G(x) = \frac{1}{M(\Theta^*)} \int_{(-\infty,x]} e^{\Theta^* y} dF(y).$$

Que tiene sentido ya que $\frac{e^{\Theta^* y}}{M(\Theta^*)} > 0$.

La definición de G es un caso muy particular del Teorema $A.2$ en el sentido de que sólo toma los conjuntos de la forma $(-\infty, x)$ que es la distibución de:

$$\tilde{G}(\cdot) = \frac{1}{M(\Theta^*)} \int_{(\cdot)} e^{\Theta^* y} dF(y).$$

Que es una probabilidad, siempre que probemos que $\tilde{G}$ es una medida de probabilidad, y esto se debe:

$$1 = Lim_{x \longrightarrow \infty} G(x) = \tilde{G}(\Omega).$$

La última igualdad es consecuencia del teorema de convergencia monótona para la sucesión de funciones crecientes y mayores que cero $\{\frac{e^{\Theta^* y}}{M(\Theta^*)} 1_{\{(-\infty,x_n]\}}\}_{n=1}^{\infty}$ con x_n una suceción creciente que diverge a ∞.

Por estar G como integral de una función integrada con la medida generada por la distribución, tenemos $G << F$, G es absolutamente continua relativa a F; por ser esta función estrictamente positiva tenemos también el inverso $F << G$, F es absolutamente continua relativa G, de hecho tenemo explicitamente lo siguiente:

$$F(x) = \int_{\{(-\infty,x]\}} 1 dF = \int_{\{(-\infty,x]\}} (\frac{1}{M(\Theta^*)} e^{\Theta^* y})^{-1} (\frac{1}{M(\Theta^*)} e^{\Theta^* y}) dF =$$
$$\int_{\{(-\infty,x]\}} (\frac{1}{M(\Theta^*)} e^{\Theta^* y})^{-1} dG$$

La primera igualdad es por escribir al 1 de forma distinta, la segunda por linealidad de la integral y la tercera es una consecuencia del teorema $A.2$.

Por lo que para cualquier α, claramente tenemos:

$$P(x \geq \alpha) = \int 1_{\{y \geq \alpha\}} dF(y) = \int 1_{\{y \geq \alpha\}} e^{\Theta^* y}(e^{\Theta^* y})^{-1} dF(y) =$$
$$M(\Theta^*) \int 1_{\{y \geq \alpha\}} (e^{-\Theta^* y}) dG(y)$$

Aplicando las ideas anteriores y usando el teorema de Fubini al lado derecho de la ecuación (3.2) tenemos,

Por Fubini

$P(X_1 + ... X_n \geq na) = \int ... \int 1_{\{Y1 + ... Y_n \geq na\}} dF(Y_1)...dF(Y_n) =$

$= \int ... \int 1_{\{Y_1 + ... Y_n \geq na\}} e^{-\Theta^*(Y_1 + ... + Y_n)} e^{\Theta^* Y_1} dF(Y_1)...e^{\Theta^* Y_n} dF(Y_n)$

Por el cambio de medida a la última ecuación,

$= M(\Theta^*)^n \int ... \int 1_{\{Y_1 + ... Y_n \geq na\}} e^{-\Theta^*(Y_1 + ... + Y_n)} dG(Y_1)...dG(Y_n) \geq$

$\geq M(\Theta^*)^n \int ... \int 1_{\{n(a+\varepsilon')Y_1 + ... Y_n \geq na\}} e^{-\Theta^*(Y_1 + ... + Y_n)} dG(Y_1)...dG(Y_n) \geq$

$M(\Theta^*)^n e^{n\Theta(a+\varepsilon')} \int ... \int 1_{\{n(a+\varepsilon') \geq Y_1 + ... Y_n \geq na\}} dG(Y_1)...dG(Y_n)$

Ahora sean, $\tilde{X}_1, ..., \tilde{X}_n$ variables aletatorias idénticamente distribuidas e independientes, con la distribución dada por G, entonces:

$$P(X_1 + ... + X_n \geq na) \geq M(\Theta^*) P(n(a + \epsilon') \geq \tilde{X}_1 + ... + \tilde{X}_n \geq na) \tag{3.5}$$

Para encontrar una cota inferior en el lado derecho de la ecuación (3.5) notamos que por ser $M(\Theta)$ finita en una vecindad abierta de Θ^* y por Teorema $A.1$ tenemos lo siguiente:

$$\frac{d^n}{d\Theta^n} M(\Theta^*) = E(X^n e^{\Theta^* X}) < \infty \text{ y esto para todo } n \in \mathbb{N}$$

En particular por el Teorema $A.2$,

$$E(\tilde{X}^n) = \frac{E(X^n e^{\Theta^* X})}{M(\Theta^*)} < \infty \text{ para todo } n \in \mathbb{N}$$

Una observación es que en Θ^* necesariamente tenemos $E((X - a)e^{\Theta^*(x-a)}) = 0$ ya que esta es la primera derivada de la función $E((X - a)e^{\Theta(X-a)})$ y en Θ^* tenemos un máximo de la función por hipótesis,

$$E(X e^{\Theta^* X}) = aM(\Theta^*)$$

Dividiento entre $M(\Theta^*)$ tenemos,

$$E(\tilde{X}) = \frac{E(X e^{\Theta^* X})}{M(\Theta^*)} = a$$

Centramos a nuestra varible aleatoria $\tilde{X}$ dada por nuestro cambio de variable exactamente en el punto a. Volviendo a la suma de las variables aleatorias $\tilde{X}_1, ..., \tilde{X}_n$ independientes e idénticamente distribuidas con todos los momentos finitos, en especial la varianza; usando el teorema del límite central, tenemos que:

$$P\Big(n(a+\epsilon') \geq \tilde{X}_1 + ... + \tilde{X}_n \geq na\Big) = P\Big(\sqrt{n}\epsilon' \geq \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\tilde{X}_i - n) \geq 0\Big).$$

Que converge a $\frac{1}{2}$. Tomando n_0 tal que la probabilidad exceda $\frac{1}{4}$ siempre que $n \geq n_0$, claramente depende de ϵ'. Entonces para $n \geq n_0$

$$P(X_1 + ...X_n \geq na) \geq \frac{1}{4}[e^{-\Theta^* a}M(\Theta^*)]^n e^{-n\Theta^*\epsilon'} = \frac{1}{4}e^{-nl(a)}e^{-n\Theta^*\epsilon'}$$

Como $\Theta^* > 0$, podemos tomar una ϵ' que satisfaga,

$\frac{1}{4}e^{-n\Theta^*\epsilon'} \geq e^{-n\Theta^*\epsilon}$ siempre que $n \geq n_0$, con esto demostramos la ecuación (3.2)

⋄

Para ver por qué la ecuación (3.2) implica la ecuación (3.3), necesitamos la siguiente proposición.

Proposición 1 *Sea $\{a_n\}$ una sucesión de números positivos que satisfacen $\forall\varepsilon > 0$ existe n_ε tal que si $n \geq n_\varepsilon$, entonces $a_n \geq e^{-n\varepsilon}$; tenemos que existe una sucesión $\{s_n\}_{n=1}^{\infty}$ que converge a cero tal que $a_n \geq e^{-ns_n}$, $a_n \geq e^{o(n)}$*

Demostración.

Notemos que con las hipótesis, Para toda $m \in \mathbb{N}$ existe $\tilde{n}_m$ tal que si $n \geq \tilde{n}_m$, entonces $a_n \geq e^{-n(\frac{1}{m})}$.

De lo que tenemos lo siguiente, Sea $n_m := max\{\tilde{n}_m, n_{m-1}\} + 1$ para $m > 1$ y para $m = 1$, $n_1 = \tilde{n}_1$.

Observación, $n_1 < n_2 < ...n_k < n_{k+1} < ...$

Definamos a $\{s_n\}_{n=1}^{\infty}$ de la siguiente manera,

para $1 < n < n_1$ $s_n = -n\ln(a_n)$,

para $n_k \leq n < n_{k+1}$ $s_n = \frac{1}{k}$, $k > 1$

por cómo construimos a nuestra sucesión tiene las siguientes dos propiedades

(i) $a_n \geq e^{-ns_n}$ y (ii) $s_n \longrightarrow 0$ cuando $n \longrightarrow \infty$. Obsérvese que $-ns_n = o(n)$.

lo que concluye la demostración.

Para ver como la demostración relaciona las ecuaciones (3.2) y (3.3), definamos a $b_n := P(X_1 + ... + X_n \geq na)$, $c_n := e^{-nl(a)}$, con lo anterior sea $a_n := \frac{b_n}{c_n}$ y por la observación $a_n \geq e^{-o(n)}$.

⋄

El siguiente ejemplo ilustra el teorema de esta sección.

Ejemplo 1 *Sean $X_1, X_2, ...$ variables aleatorias con distribución normal con esperanza igual a cero y varianza igual a uno.*

Entonces $M(\Theta) = e^{\frac{1}{2}\Theta^2}$, $l(a) = sup_\Theta \, a\Theta - \frac{1}{2}\Theta^2$ Derivando e igualando a cero tenemos,

$a = \Theta$, es una función continua, creciente y derivable ponicdo a a en función de Θ y viceversa. Lo anterior se debe a algo más general que se vera en la proposición 7 de la siguiente sección.

$l(a) = \frac{1}{2}a^2$.

Ahora como sabemos la suma de variables aleatorias $X_1, ..., X_n$ independientes y con distribución normal con paramentros (μ_i, σ_i^2), es una normal con parametros $(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$, en nuestro caso $X_1 \sim N(0,1)$ y $\sum_{i=1}^n X_i \sim N(0,n)$

$$\begin{aligned} P(X_1 + ... + X_n \geq na) &= \int_{na}^{\infty} \frac{e^{\frac{-Y^2}{2n}}}{\sqrt{2\pi n}} dY = \\ &= \int_{\sqrt{n}a}^{\infty} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt \end{aligned} \tag{3.6}$$

Una estimación para la integral (3.6) la podemos encontrar en el Feller, An Introduction to Probability and Its Applications, página 175:

$$[X^{-1} + X^{-3}]\tfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}X^2} < \textstyle\int_X^\infty \tfrac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}dt < X^{-1}\tfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}X^2}$$

en nuestro caso,

$$[(\sqrt{n}a)^{-1} + (\sqrt{n}a)^{-3}]\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\sqrt{n}a)^2} < \int_{\sqrt{n}a}^{\infty} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}dt < (\sqrt{n}a)^{-1}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\sqrt{n}a)^2} \tag{3.7}$$

Que nos da,

$$\int_{\sqrt{n}a}^{\infty} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}}dt \sim \frac{1}{\sqrt{2n\pi}a}e^{-\frac{1}{2}n(a)^2} \tag{3.8}$$

Que concuerda con el teorema.

Capítulo 4

Observaciones de la Función Tasa

4.1. Función Tasa

En este capítulo seguiremos discutiendo el comportamiento asintótico de

$$a_n = P\Big(\frac{X_1 + ... + X_n}{n} \geq a\Big) \tag{4.1}$$

donde las X_i son variables aleatorias independientes e idénticamente distribuidas con $E(X_i) < a$. Por la ley débil de los grandes números sabemos que la sucesión anterior converge a cero, conforme n tiende a infinito. El teorema del capítulo anterior nos dice cómo converge la sucesión $\{a_n\}_{n=1}^{\infty}$,

$$a_n = \exp(-nl(a) + o(n)) \tag{4.2}$$

donde

$$l(a) = -\ln(inf_{\Theta\in\Re}\{\exp(-\Theta a)M(\Theta)\}) = sup_{\Theta\in\Re}(a\Theta - \ln(M(\Theta))) \tag{4.3}$$

En esta sección nuestro primer objetivo es entender mejor a $l(a)$ a partir de las siguientes propiedades de la función

$$f(\Theta) = a\Theta - \ln M(\Theta) \tag{4.4}$$

1. $f(\Theta)$ es una función cóncava.
2. $f^{'}(0) > 0$.
3. Encontrar una cota para el supremo de la función $f(\Theta)$.
4. Discutir en qué casos basta fijarnos en $\Theta \geq 0$ para encontrar el supremo de $f(\Theta)$.

5. Daremos condiciones necesarias y suficientes para que el supremo sea un máximo.

De las cinco observaciones anteriores unas primeras conclusiones son: Un máximo local de $f(\Theta)$ es un máximo global, basta encontrar una raíz de la derivada para conocer dónde se alcanza el máximo, pasamos de un problema de maximizar a otro de buscar raíces. También daremos condiciones para que se satisfagan las propiedades:

6. Encontraremos un intervalo abierto en el que $l(\cdot)$ es continua.

7. Con el intervalo abierto anterior tendremos una biyección entre las a y las Θ del máximo de la ecuación (4.4), esta biyección resultara ser una función diferenciable con inversa diferenciable. De hecho existe una relación entre la continuidad de $l(\cdot)$ en a y que la ecuación (3) sea un máximo.

Sabremos precisamente el intervalo en el que $l(\cdot)$ es continua. Éste dependerá de nuestras variables aleatorias.

La parte teórica de esta sección vendrá complementada con ejemplos que ilustran los distintos problemas y comportamientos de la función $l(\cdot)$.

El entender mejor la función $f(\Theta)$ permitirá comprender mejor la hipótesis de Θ^* del teorema del capítulo anterior, nuestra primera extensión de este teorema eliminará esta hipótesis, la segunda extensión buscará generalizar el teorema a otro tipo de conjuntos, en ambos casos requerimos de la continuidad l($\cdot$) de 6 . Por último demostraremos una versión de la ley fuerte de los grandes números usando la teoría de esta sección.

Proposición 2 *Sea X una variable aleatoria tal que $f(\Theta)$ está definida en D con $0 \in D^o$ entonces se satisface que f es una función cóncava.*

Demostración.

Por el teorema A.1, $E(\exp(\Theta X))$ tiene primera y segunda derivada, de lo que se sigue que $f(\Theta)$ también. Ahora calculando la primera derivada tenemos:

$$f'(\Theta) = a - \frac{E(X\exp(\Theta X))}{E(\exp(\Theta X))}. \tag{4.5}$$

De lo anterior tenemos que la segunda derivada es:

$$f''(\Theta) = \frac{-E(X^2e^{\Theta X})E(e^{\Theta X}) + E^2(Xe^{\Theta X})}{E^2(e^{\Theta X})}. \tag{4.6}$$

Ahora bien, usando el cambio de medida del teorema A.2 con la función $g = e^{\Theta X}$ tenemos que la ecuación anterior se reduce a

$$f''(\Theta) = \frac{-E_g(X^2) + E_g^2(X)}{1}. \tag{4.7}$$

Pero lo anterior es una $-Var_g(X)$ que es siempre menor o igual a cero.

Más adelante, para saber algunos casos en que $l(a)$ es una función continua nos interesará saber cuándo $f^{''}(\Theta)$ es menor que cero, véase proposición 9. La siguiente observación es una consecuencia de la ecuación (4.5) de nuestra proposición 1.

Observación 1 *Sea X una variable aleatoria tal que $f(\Theta)$ está definida en D con $0 \in D^o$ y $a > E(X)$, entonces $f^{'}(0) > 0$.*

Demostración
Evaluando en 0 la ecuación (4.5) tenemos

$$f^{'}(0) = a - E(X) \tag{4.8}$$

que por hipótesis es mayor o igual que 0.

La siguiente proposición es una consecuencia de la desigualdad de Chernoff; una demostración de ella que requiere del uso del cambio de variable, como la del teorema A.2 además de probabilidad condicional, se puede consultar en el Ross, A second course in Probability, pág 119, aquí presentamos una demostración alternativa.

Proposición 3 *Sea x una variable aleatoria tal que $f(\Theta)$ está definida en D con $0 \in D^o$, entonces tenemos que $f(\Theta) < -\ln(P(X \geq a)) \quad \forall\Theta > 0$.*

Demostración.

Notando que $f(\Theta)$ es lo mismo que $-\ln(e^{-\Theta a}M(\Theta))$, vamos a encontrar cotas para la ecuación dentro de la función logaritmo:

$$\begin{aligned} e^{-\Theta a}M(\Theta) &= E(e^{\Theta(X-a)}) = \int e^{\Theta(X-a)}dP = \\ &= \int_{\{X-a\geq 0\}} e^{\Theta(X-a)}dP + \int_{\{X-a<0\}} e^{\Theta(X-a)}dP. \end{aligned}$$

Ahora la integral del lado derecho es mayor o igual que cero; por lo que lo anterior es mayor o igual que:

$$\geq \int_{\{X-a\geq 0\}} e^{\Theta(X-a)\}}dP.$$

Notando que en $\{X - a \geq 0\} = \{X \geq a\}$ y $\Theta > 0$ la función que integramos es mayor o igual a 1 tenemos:

$$\geq \int_{\{X-a\geq 0\}} e^{\Theta(X-a)}dP \geq \int_{\{X-a\geq 0\}} 1dP = = P(X - a \geq 0) = P(X \geq a).$$

Por lo que podemos concluir que $e^{-\Theta a}M(\Theta) \geq P(X \geq a)$. Que a su vez se reduce a que $\ln(e^{\Theta a}M(\Theta)) \geq \ln(P(X \geq a))$.

Multiplicando por menos y tomando ínfimo sobre $\Theta > 0$ en la última desigualdad podemos decir lo siguiente:

$$l(a) \leq -\ln(P(X \geq a)) = \ln(\frac{1}{P(X \geq a)}) \tag{4.9}$$

entre más chica sea a, tendremos una cota menor.

$\diamond$

Para ver cómo esta cota no se puede mejorar, véase ejemplo 2.

Proposición 4 *Sea X una variable aleatoria tal que su función generadora de momentos $M(\Theta) : D \longrightarrow \Re$ satisface $0 \in D^o$, $E(X) = 0$; entonces tenemos que $f(\Theta) \leq 1$ si $\Theta \leq 0$ y se satisface con igualdad cuando $\Theta = 0$.*

Demostración.

Por la desigualdad de Jensen, Proposición A.2

$$M(\Theta) = E(e^{\Theta X}) \geq e^{E(\Theta X)} = e^{\Theta E(X)} = e^0 = 1.$$

Y para una $a > E(X) = 0$ y $\Theta \leq 0$.

$$e^{-\Theta a} M(\Theta) \geq 1 \text{ y se da con igualdad para } \Theta = 0.$$

Aplicando logaritmo de ambos lados y luego multiplicando por -1

$$a\Theta - \ln\left(M(\Theta)\right) \leq 1.$$

$\diamond$

Una pequeña observación que permite centrar la esperanza en cero, es que si X es una variable y defino a $Y = X + b$ con $b \in \Re$, la relación de su funciones generadoras de momentos y sus respectivas transformadas de Cramér l_x, l_y está dada por

$M_y(\Theta) = E(e^{\Theta(X+b)}) = E(e^{\Theta X})e^{\Theta b} = M_x(\Theta)e^{\Theta b}$ y de la relación,
$a\Theta - \ln(M_y(\Theta)) = a\Theta - \ln(M_x(\Theta)) - (\Theta b) = (a-b)\Theta - \ln(M_x(\Theta))$, tenemos que:

$$M_y(\Theta) = M_x(\Theta)e^{\Theta b} \tag{4.10}$$

$$l_y(a) = l_x(a-b). \tag{4.11}$$

Proposición 5 *Supongamos que $E(\exp(\Theta X))$ está definida para todo $\Theta \geq 0$ y $P(X > a) > 0$; entonces tenemos lo siguiente:*

$$lim_{\Theta \longrightarrow \infty} f^{'}(\Theta) < 0.$$

Demostración

Como sólo nos interesa el signo de la primera derivada y con la ecuación (4.5) podemos sólo preocuparnos por el signo de $aE(e^{\Theta X}) - E(xe^{\Theta X})$. Usando la definición de esperanza obtenemos lo siguiente:

$$aE(e^{\Theta X}) - E(xe^{\Theta X}) = a\int e^{\Theta X}dP - \int Xe^{\Theta X}dP =$$

De la linealidad de la integral se sigue:

$$= \int ae^{\Theta X} - Xe^{\Theta X}dP = \int (a-X)e^{\Theta X}dP =$$

Dado que $\{a \geq X\}\bigcap\{a < X\} = \emptyset$ tenemos,

$$= \int_{\{a\geq X\}}(a-X)e^{\Theta X}dP + \int_{\{a<X\}}(a-X)e^{\Theta X}dP \leq$$

De la monotonía de la integral.

$$\leq \int_{\{a\geq X\}}(a-X)e^{\Theta a}dP + \int_{\{a<X\}}(a-X)e^{\Theta X}dP =$$
$$= e^{\Theta a}\int_{\{a\geq X\}}(a-X)dP + \int_{\{a<X\}}(a-X)e^{\Theta X}dP \leq (*)$$

Dado que $(a-X)e^{\Theta X}1_{\{a<X\}} < 0 \quad$ y $\quad 1_{\{a+\varepsilon<X\}} \leq 1_{\{a<X\}}$ se sigue que $(a-X)e^{\Theta X}1_{\{a<X\}} \leq (a-X)e^{\Theta X}1_{\{a+\varepsilon<X\}}$.
Aquí se requiere que $P(a < X) > 0$, de hecho la observación 2 muestra cómo no es suficiente que $P(a \leq X) > 0$

$$(*) \leq e^{\Theta a}\int_{\{a\geq X\}}(a-X)dP + \int_{\{a+\varepsilon<X\}}(a-X)e^{\Theta X}dP \leq$$

Por el conjunto donde integramos el lado derecho $\{a+\varepsilon < X\}$ tenemos que $a - X < -\varepsilon$

$$\leq e^{\Theta a}\int_{\{a\geq X\}}(a-X)dP + \int_{\{a<X+\varepsilon\}}(-\varepsilon)e^{\Theta X}dP =$$

Por la misma razón, en $\{a+\varepsilon < X\}$ y $\Theta \geq 0$ ocurre que $e^{\Theta a+\varepsilon} < e^{\Theta X}$ y la desigualdad se invierte al multiplicar por $-\varepsilon$, de lo que:

$$= e^{\Theta a}\int_{\{a\geq X\}}(a-X)dP + e^{\Theta(a+\varepsilon)}\int_{\{a<X+\varepsilon\}}(-\varepsilon)dP$$

Definiendo a $c_1 := \int_{\{a\geq X\}}(a-X)dP$ y a $c_2 := \int_{\{a<X+\varepsilon\}}(\varepsilon)dP$ tenemos:

$$= c_1 e^{\Theta a} - c_2 e^{(\Theta a+\varepsilon)}$$

Y por propiedades de la exponencial existe una Θ^* tal que si $\Theta \geq \Theta^*$ entonces la última formula es negativa.

$\diamond$

Observación 2 *Como se puede observar de la demostración, el caso en el que $P(a < X) = 0$ implica que $f^{'}(\Theta) > 0$ se sigue de:*

$aE(e^{\Theta X}) - E(Xe^{\Theta X}) = \int_{\{a \geq X\}}(a - X)e^{\Theta X}dP + \int_{\{a<X\}}(a - X)e^{\Theta X}dP =$
$= \int_{\{a \geq X\}}(a - X)e^{\Theta X}dP$
que es siempre estrictamente mayor que cero, eliminando el caso en el que $P(X = a) = 1$, véase ejemplo 1.

La cota dada por la proposición 3, nos dice que aunque ocurra

$P(a < X) = 0$, en caso de que $P(X = a) > 0$ se tiene:

$$Lim_{\Theta \longrightarrow \infty} f^{'} = 0.$$

Cuando tengamos $P(a \leq X) = 0$, observamos lo siguiente:

$$a\Theta - \ln(Ee^{\Theta(X-a)}) = -\ln(\int_{\{X-a<0\}} e^{\Theta(X-a)})$$

Pero ocurre que $|e^{\Theta(X-a)}1_{X-a<0}| \leq 1$, $Lim_{\Theta \longrightarrow \infty} e^{\Theta(X-a)} = 0$ y 1 es integrable. Del teorema de convergencia dominada tenemos que:

$$Lim_{\Theta \longrightarrow \infty} \int_{\{X-a<0\}} e^{\Theta(X-a)} = 0.$$

Y por continuidad de la función logaritmo, tenemos:

$$Lim_{\Theta \longrightarrow \infty} - \ln(\int_{\{X-a<0\}} e^{\Theta(X-a)}) \text{ diverge a infinito.}$$

$\diamond$

La siguiente conclusión tiene implicaciones al tratar de buscar condiciones para la continuidad de $l(\cdot)$, véase proposición 8 y los ejemplos 1, 2.

Conclusión 1 *Si $P(a \leq X) = 0$, entonces*

$$l(a) = Lim_{\Theta \longrightarrow \infty} - \ln(\int_{\{X-a<0\}} e^{\Theta(X-a)}) = \infty.$$

Proposición 6 *Sea x una variable aleatoria y a tal que $a > E(X)$ y tal que su función generadora de momentos, $M(\Theta) : D \longrightarrow \Re$, satisfaga que $0 \in D^o$ y $b := sup\, D < \infty$.*

Entonces $Lim_{\Theta \longrightarrow b^-} a\Theta - \ln(E(e^{\Theta X}))$ tampoco existe.

La demostración se hará por contradicción.

Demostración.

Supongamos que:

$$Lim_{\Theta \longrightarrow b^-} a\Theta - \ln(E(e^{\Theta X})) = c.$$

Ahora sabemos que Θ,a,$\frac{1}{\Theta}$ convergen cuando Θ tiene a b por la derecha, y como producto, suma de sucesiones convergentes vuelve a ser convergente tenemos:

$$Lim_{\Theta \longrightarrow b^-} [\frac{a\Theta - \ln(E(e^{\Theta X}))}{\Theta} - a]\Theta = b(\frac{c}{b} - a)$$

con lo que concluimos:

$$Lim_{\Theta \longrightarrow b^-} - \ln(E(e^{\Theta X})) = b(\frac{c}{b} - a)$$

que por continudad de logaritmo obtenemos:

$$Lim_{\Theta \longrightarrow b^-} (E(e^{\Theta X})) = e^{b(\frac{c}{b} - a)}.$$

Es decir existe el limite de

$$Lim_{\Theta \longrightarrow b^-} E(e^{\Theta X}).$$

¡Contradicción!

$\diamond$

Veamos por qué el límite no puede ser de la forma $\{(-1)^n\}_{n=1}^{\infty}$, es decir oscilante, sino que diverge a menos infinito. Para ello ocupamos el siguiente resultado:

Teorema 1 *Sea $\varphi : D \longrightarrow \Re$ una función convexa, entonces $\forall x \in D^o$ $\varphi(x)$ es continua y tiene derivada derecha e izquierda, no necesariamente iguales, en particular φ es continua en D^o*

Su demostración se encuentra en un teorema más general, véase Teorema A.5 .

$\diamond$

Proposición 7 *Sea X una variable aleatoria y a tal que $a > E(X)$ y tal que su función generadora, $M(\Theta) : D \longrightarrow \Re$, satisfaga que $0 \in D^o$ y y $b := sup\, D < \infty$*

entonces $a\Theta - \ln(E(e^{\Theta X}))$ diverge cuando $\Theta \longrightarrow b^-$, de hecho, diverge a $-\infty$, si $P(X \geq a) > 0$.

Demostración.

Del teorema anterior tenemos que $a\Theta - \ln(E(e^{\Theta X}))$ diverge o converge cuando $\Theta \longrightarrow b^-$, pero no puede converger por la proposición 6

Por lo tanto, tenemos por la proposición 2. una cota superior,

$-\ln(P(X \geq a))$, por lo que el límite sólo puede ser $-\infty$.

$\diamond$

Puede verificarse fácilmente que el suponer que $P(X \geq a) > 0$ no es realmente una restricción adicional, sino es consecuencia de las tres hipótesis de la observación anterior, véase observación 3.

Teorema 2 *Sea X una variable aleatoria que satisface $P(X \in (-\infty, t)) = 1$ para alguna $t \in \Re$, entonces su función generadora de momentos está definida en todos los reales mayores o iguales que cero.*

Demostración.

El caso $\Theta = 0$, es trivial.

Sea $\Theta > 0$

$$M(\Theta) = E(e^{\Theta X}) = \int_{\{X\leq 0\}} e^{\Theta X} dP + \int_{\{X>0\}} e^{\Theta X} dP \leq$$
$$\int_{\{X\leq 0\}} 1 dP + \int_{\{X>0\}} e^{\Theta X} dP =$$

$$= P(X \leq 0) + \int_{\{X>0\}} e^{\Theta X} dP \leq P(X \leq 0) + e^{\Theta t} \int_{\{X>0\}} dP =$$

$$= P(X \leq 0) + e^{\Theta t} P(X > 0)$$

Que es es un real para toda $\Theta > 0$.

$\diamond$

Por contrapositiva tenemos que:

Observación 3 *Sea x una variable aleatoria y a tal que $a > E(X)$ y tal que su función generadora, $M(\Theta) : D \longrightarrow \Re$, satisfaga que $0 \in D^o$ y $b := sup\, D < \infty$, entonces $\forall a > 0$ $P(X \leq a) > 0$*

Observación 4 *Sea x una variable aleatoria tal que su función generadora, $M(\Theta) : D \longrightarrow \Re$, satisfaga:*

i) $0 \in D^o$ ii) $b := sup\, D$, b puede ser infinito; entonces ocurre lo siguiente:

Si $P(X > a) > 0$, entonces $l(a) = max_{\Theta \in (0,b)} f(\Theta)$.

Si $P(X > a) = 0$, entonces $l(a) = Lim_{\Theta \longrightarrow \infty} f(\Theta)$.

Observación 5 *En el caso de que $P(X > a) > 0$ por proposiciones 5 y la observación anterior siempre podemos encontrar un intervalo abierto U que satisface $\Theta^* \subset U$ y $U \subset (0, b)$, en el caso en el que la función generadora de momentos esté definida para todo real positivo, tomamos b tal que $f'(b) < 0$. Esto será útil en la demostración de la proposición 8, sobre la continuidad de $l(a)$*

Para variables aleatorias donde la función generadora de momentos está definida en todos los reales e inclusive con sólo estarlo en una vecindad del cero, tenemos más aún, de las proposiciones 2, 5 y 7 que podemos garantizar que para cualquier a con $a > E(X)$ y $P(X \geq a) > 0$, obtenemos $l(a) = f(\Theta^*)$ para alguna Θ^*. Esto último y con la condición de que exista $b > a$ tal que $P(X \geq b) > 0$ nos permitirá obtener continuidad sobre $l(a)$ $a \in (E(X), \infty)$.

Los siguientes dos teoremas con su demostración se encuentran en el Burkill and Burkill, A second course in Mathematical Analysis, páginas 203 y 212, y su uso sólo se limita a esta sección.

Teorema 3 *Sea $A \subseteq \Re^m$ y sea u un punto interior de A. Si $f : A \longrightarrow \Re^n$ es tal que cada derivada parcial $D_i f_j$ $(j = 1, ...n; i = 1, ...m)$ existen en una bola abierta $B_\delta(u)$ $\delta > 0$ y es continua en u, entonces f es diferenciable en u*

Teorema 4 (TEOREMA DE LA FUNCIÓN IMPLÍCITA)

Sea (u, v) un punto interior de $E \subsetneq \Re^2$ y supongamos que la función

$f : E \longrightarrow \Re$ satisface las siguientes condiciones:

(i)$f(u, v) = 0$

(ii) f es continuamente diferenciable en conjunto G abierto y que contiene al punto (u, v)

(iii)$D_2 f(u, v) \neq 0$

Entonces existen un rectángulo

$$M \times N = [u - \varepsilon, u + \varepsilon] \times [v - \delta, v + \delta]$$

y una función continua $\phi : M \longrightarrow N$ tal que $y = \phi(x)$ y es la única solución en $M \times N$ de la ecuación $f(x, y) = 0$; más aún ϕ es continuamente diferenciable en M° y

$$\phi'(x) = \frac{-Df_1(x,\phi(x))}{Df_2(x,\phi(x))}$$

para toda $u - \varepsilon < x < u + \varepsilon$

Proposición 8 *Sea X una variable aleatoria y a tal que $E(X) < a < b$ donde $P(X \geq b) > 0$, $l(a) = f(\Theta^*)$ para alguna Θ^*; entonces $l(\cdot)$ es continua en a.*

La demostración nos dará las siguientes observaciones.

Observación 6 *Existe $\phi : D \longrightarrow I$, con $a \in D$ y $\Theta^* \in I$ función biyectiva diferenciable, con inversa diferenciable y que pone en correspondencia uno a uno entre las a y el óptimo de la función $f(\Theta) = a\Theta - \ln(M(\Theta))$.*

Observación 7 *Sea X una variable aleatoria con función generadora de momentos definida en un intervalo abierto del cero.*

Si $P(a < X) > 0$ para toda $a > E(X)$, entonces $l(\cdot) : (E(X), \infty) \longrightarrow \Re$ es continua.

Si $P(a < X) = 0$ para algún $a > E(X)$, entonces $l(\cdot) : (E(X), b) \longrightarrow \Re$ es continua, donde $b := inf\{a \in (E(X), \infty) : P(X > a) = 0\}$

Demostración.

Una primera aclaración, la función f con la que hemos estado trabajando y que está definida hasta ahora en $\Re$ la podemos extenderla a $\Re^2$ haciendo a a variable, $f(\Theta, a)$ Recordando que $l(a)$ es el supremo de la función $f(\cdot)$ que como ya se vio es un máximo en el caso que $P(x > a) > 0$ y se alcanza en la única raíz de la función $g(\Theta, a) = a - \frac{E(Xe^{\Theta}X)}{E(e^{\Theta}X)}$. Ahora notamos que $g_a = 1$ y es continua en todo $\Re$, sea B un intervalo abierto que contiene a a.

También, $g_\Theta = f''$ que es continua en el intervalo $\Theta^* \in U$, de la observación 5. Por el teorema 3, existe $B_\delta := B_{\delta'}(\Theta^*, a) \subset U \times B$ tal que $g : B_\delta \longrightarrow \Re$ es diferenciable, además $g(\Theta^*, a) = 0$ y $g_\Theta \neq 0$.

Por lo que podemos concluir que existen $\epsilon, \delta > 0$ y una función $\phi : [a - \epsilon, a + \epsilon] \longrightarrow [\Theta^* - \delta, \Theta^* + \delta]$ tal que $y = \phi(x) = \Theta(x)$ y es la única solución a $f_\prime(y, x) = 0$ pero entonces $l(a) = f(\Theta(a), a)$ que es composición de funciones continuas.

De lo anterior tenemos una relación local de $l(a)$ y las Θ^*, el supremo es un máximo y está dada por la función ϕ. Por último notemos que:

$$\acute{l}(a) = (f(\Theta(a), a))' = f_1(\Theta(x), a)\Theta'(x) + f_2(a\Theta(a), a) \tag{4.12}$$

$\diamond$

Para terminar la demostración anterior basta ver por qué la segunda derivada de f no es cero.

Proposición 9 *Sea x una variable aleatoria, entonces $E((X - E(X))^2) \neq 0$ si y sólamente si bajo cualquier cambio de medida dada por el Teorema A.2, sea g, $E_g((X - E_g(X))^2) \neq 0$*

Demostración

$E_g((X - E_g(X))^2) = E((X - E(X))^2 g) \neq 0$ si y sólamente si $(X - E(X))^2 g = 0$ casi seguramente, pero $g > 0$, vease A.2, por lo que $(X - E(X))^2 g = 0$ c.d si y sólo si $(X - E(X))^2 \neq 0$ c.d que a su vez se cumple si y sólo si $E((X - E(X))^2) \neq 0$

Es decir $Var_g(X) \neq 0$ si y solamente si $Var(X) \neq 0$

$\diamond$

De la proposición 5 y las observaciones 2 y 4, las hipótesis de la proposición 7 pueden sustituirse por las siguientes.

I. **Sea X una variable aleatoria tal que su función generadora de momentos $M(\Theta)$ esté definida en D con $0 \in D^o$ y tal que $l(a) = f(\Theta^*)$**

II. **Sea x una variable aleatoria tal que su función generadora de momentos $M(\Theta)$ esté definida en D con $0 \in D^*$ y tal que $P(a < X) > 0$**

La conclusión 1 nos dice que $l(a)$ no puede ser continua en los puntos en los que $P(X = a) > 0$ y $P(X > a) = 0$; véase, ejemplo 2, la Bernoulli con $a = 1$.

Uno espera que entre más alejado se esté de la esperanza, más rápido se va a cero, esta idea intuitiva se confirma con la siguiente observación:

La función, $l(\cdot)$, de la proposición 8 es creciente por la siguiente razón, sea $a < a'$, y Θ^{*_a} tal que $aE(e^{\Theta^{*_a}X}) - E(e^{\Theta^{*_a}X}) = 0$, es la raíz de la derivada de la función: $a - \ln(M(\Theta))$, entonces como $E(e^{\Theta^{*_a}X}) > 0$, se sigue que $a'E(e^{\Theta^{*_a}X}) - E(e^{\Theta^{*_a}X}) > aE(e^{\Theta^{*_a}X}) - E(e^{\Theta^{*_a}X}) = 0$, y por concavidad de $a'\Theta - \ln(M(\Theta))$ el máximo se alcanza en alguna $\Theta > \Theta^{*_a}$, la desigualdad es estricta. Es decir, si $P(a' < X)$, entonces $l(a) < l(a')$.

La convexidad de $l(\cdot)$ se sigue de lo siguiente,

Sea $\lambda \in (0, 1)$ y $E(X) < a < a'$ y $P(a < X) > 0$,

$$\begin{aligned}\lambda l(a) + (1-\lambda)l(a') &= \lambda sup_{\Theta>0}\{a\Theta - M(\Theta)\} + (1-\lambda)sup_{\Theta>0}\{a'\Theta - M(\Theta)\} = \\ &= sup_{\Theta>0}\{\lambda a\Theta - \lambda M(\Theta)\} + sup_{\Theta>0}\{(1-\lambda)a'\Theta - (1-\lambda)M(\Theta)\} \geq \\ &\quad sup_{\Theta>0}\{(\lambda a + (1-\lambda)a')\Theta - M(\Theta)\} = l(\lambda a + (1-\lambda)a').\end{aligned}$$

Notemos, cuando $P(a \leq X) = 0$, tenemos que $l(\cdot) : [a, \infty) \longrightarrow \Re$ es creciente y convexa, todo se reduce a infinitos.

La siguiente observación es consecuencia de la discusión anterior,

Observación 8 *La función $l(\cdot) : (E(X), \infty) \longrightarrow \Re$ es una función creciente y convexa.*

Queda un pequeño detalle en la observación 8 y es el caso cuando $P(X = a)$ y $P(a < X) = 0$, este detalle quedará completo con las siguientes dos proposiciones. Estos permitirán extender la continuidad de la proposición 8 y el teorema del capítulo anterior para cuando no esté garantizada la Θ^* de la hipótesis.

Proposición 10 *Sea X una variable aleatoria tal que su función generadora de momentos esté definida en una vecindad del cero $b > E(X)$ es tal que $P(X = b) > 0$ y $P(X > b) = 0$.*

Si $\{b_m\}_{m=1}^{\infty} \subset (E(X), b)$ con $b_m \uparrow b$, entonces $l(b_m) \uparrow l(b)$.

Demostración. Sea $\varepsilon > 0$

Por la observación 2 sabemos que

$$l(b) = Lim_{\Theta \longrightarrow \infty} b\Theta - \ln(M(\Theta)$$

como es finito y es el supremos de la función $f(\Theta)$ tomemos $\bar{\Theta}$ tal $|b\bar{\Theta} - \ln(M(\bar{\Theta}) - l(b)| < \frac{\varepsilon}{2}$.

Ahora sea m tal que si $n \geq m$, $|b_n - b|\bar{\Theta} < \frac{\varepsilon}{2}$, de lo anterior se sigue.

$$|b_n\bar{\Theta} - \ln(M(\bar{\Theta})) - l(b)| \leq$$
$$|b_n\bar{\Theta} - \ln(M(\bar{\Theta})) - [b\bar{\Theta} - \ln(M(\bar{\Theta}))]| + |b\bar{\Theta} - \ln(M(\bar{\Theta})) - l(b)| =$$
$$= |b_n - b|\bar{\Theta} + |b\bar{\Theta} - \ln(M(\bar{\Theta})) - l(b)| \leq \varepsilon.$$

Pero $b_n\bar{\Theta} - \ln(M(\bar{\Theta})) \leq l(b_n) \leq l(b)$. que nos da $|l(b_n) - l(b)| < \varepsilon$ cuando $n \geq m$.

$\diamond$

Observación 9 *Sea X una variable aleatoria, con $Var(X) > 0$. con función generadora de momentos definida en un intervalo abierto alrededor del cero.*

Si X es tal que $P(X > a) > 0$ para toda $a > E(X)$, entonces

$l(\cdot) : (E(X), \infty) \longrightarrow \Re$ es continua.

Si X es tal que $P(X > a) = 0$ con $P(X = a) > 0$, entonces

$l(\cdot) : (E(X), a] \longrightarrow \Re$ es continua.

El siguiente resultado extiende el teorema del capítulo 1.

Proposición 11 *Consideremos una sucesión $\{X_i\}_{i=1}^{\infty}$ de variables aleatorias independientes e idénticamente distribuidas, con función generadora de momentos definida en un intervalo abierto alrededor del cero. Para cualquier $a > E(X)$ tenemos*

$$Lim_{n \longrightarrow \infty} \frac{1}{n} \ln\left(P(X_1 + ...X_n \geq na)\right) = l(a).$$

Por el capítulo 1 sólo nos basta probar el caso en el que $P(X = a) > 0$ y $P(X > a) = 0$. En la demostración requerimos el siguiente teorema sobre sucesiones dobles y que se puede consultar en: Robert G. Bartle, Una Introducción al Análisis Matemático, pág 156.

Teorema 5 *Sea $x_{m\,n}$ una sucesión doble y $x \in \Re$ tal que $\forall\, \varepsilon > 0$ existe N_ε tal que si $n, m \geq N_\varepsilon$,*

$|x - x_{m\,n}| < \varepsilon$, es decir que la sucesión doble converge a x, y que los límites $Lim_{n \longrightarrow \infty} x_{m\,n} = y_m$, $Lim_{m \longrightarrow \infty} x_{m\,n} = z_m$ existen para todos los números naturales m, n. Entonces los límites iterados

que se reduce a

$$P(Lim_{n\longrightarrow\infty}Z(n,x)=0)=1 \tag{5.11}$$

Para la demostración de (iii), consideremos una n^* con la propiedad de que,

$$E(Z(n,x))=b^{n^*}P(w_1+...+w_n\geq 0)>1$$

Esto se puede porque $P(w_1+...+w_n\geq 0)=e^{\ln(b)+l(0)}$. Construyamos un proceso de ramificación $K(k)$ de la siguiente manera. Los descendientes en la primera generación son los miembros del proceso de ramificación aleatorio que comenzando en cero, $x_0=0$, son positivos en la generación n^*, así las cosas, $K(1)=Z(n^*,0)$, de este modo la distribución del proceso de ramificación para $K(k)$ tiene distribución $Z(n^*,0)$. Así, hemos colocado la esperanza de este proceso de ramificación de tal manera que es mayor estricto que 1. Por el teorema 12 sabemos que la probabilidad de que $K(k)\longrightarrow\infty$ conforme k tiende a infinito es mayor estricto que cero. Ahora, construyamos K de tal modo que $K(k)\leq Z(kn^*,0)$, esto se logra ignorando de $x(kn^*,i)$ los valores negativos y adicionalmete los que son menores que el valor de alguno de sus ancestros, que es lo mismo a pedir que sea mayor que el de su progenitor, podemos sustituir $x_0=0$ por $x_0=y$ y la palabra positivos por mayores que y. Esto garantiza que $K(k)\leq Z(kn^*,0)$.

Como $Z(n,x)\geq Z(n,0)$ para toda $x\geq 0$, podemos concluir que la probabilidad de que $Z(n,x)\longrightarrow\infty$ a lo largo de la sucesión kn^* es positiva

Obsérvese cómo el resultado anterior sólo depende de b, no de la distribución del número de hijos. Sin embargo, para el cálculo de $l(0)$ se toma en cuenta qué tan alejado estoy de $E(w_n)$ y la distribución de w_n. El siguiente resultado resulta inmediato,

Corolario 2 *El teorema de Biggins y el lema 2 siguen siendo válidos si sustituimos $Z(n,x)$ por el número de individuos de la generación n que son estrictamente positivos.*

Para el Lema 2 obsérvese que la propiedad que estamos sustituyendo en este corolario no se utilizó en la demostración. Y para el teorema de Biggins hay que recordar que el teorema principal es válido para conjuntos de la forma (a,∞), en particular para $x=0$.

Una variación interesante la tenemos cuando nuestra población tiene una cierta habilidad (por ejemplo la fuerza, la velocidad con que corre, etc) y los únicos individuos que se reproducen son los que el valor de esta habilidad no es negativa. El número de hijos de un individuo que se puede reproducir no depende del valor de su habilidad. Queda claro que la extinción de los descendientes de un individuo no sólo depende de la esperanza del número de hijos, también que a lo largo de las generaciones siempre existe alguien que

$$Lim_{n\longrightarrow\infty}\, Lim_{m\longrightarrow\infty} x_{m\,n} \quad y \quad Lim_{m\longrightarrow\infty}\,,Lim_{n\longrightarrow\infty} x_{m\,n}$$

existen y son iguales a x

Demostración de la proposición 11.

Sea $\{a_m\}_{m=1}^{\infty} \subset (E(X), a)$ una sucesión creciente que converge a a. Definamos a $c_{m\,n} = \frac{1}{n}\ln\left(P(X_1 + ...X_n \geq na_m)\right)$.

Observación 1.

$$Lim_{m\longrightarrow\infty} c_{m\,n} = \tfrac{1}{n}\ln\left(P(X_1 + ...X_n \geq na)\right).$$

Observación 2.

$$Lim_{n\longrightarrow\infty} c_{m\,n} = Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\left(P(X_1 + ...X_n \geq na_m)\right) = l(a_m).$$

Este último limite es válido porque $a_m \subset (E(X), a)$ y podemos garantizar la existencia de Θ^* de la hipótesis del teorema del capítulo 1 Para la observación 3., que la sucesión doble converge a $l(a)$, sea $\varepsilon > 0$.

Por la proposición 10 podemos tomar m_ε tal que si $m \geq m_\varepsilon \quad l(a_m) - l(a)| < \frac{\varepsilon}{2}$.

Ahora tomemos n_ε tal que si $n \geq n_\varepsilon$
$|\frac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_{m_\varepsilon})\right) - l(a_{m_\varepsilon})| < \frac{\varepsilon}{2}$, esto se puede porque dejamos fija a a_{m_ε}.

Nótese que: $|\frac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_{m_\varepsilon})\right) - l(a)| < \varepsilon$, por desigualdad del triángulo.

Ahora sea $m \geq m_\varepsilon$ y $n \geq n_\varepsilon$

$$|\tfrac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_m)\right) - l(a)| <$$
$$|\tfrac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_{m_\varepsilon})\right) - l(a)| < \varepsilon.$$

La primera desigualdad se debe a que $\frac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_m\right)$ decrece conforme m crece y $-l(a_m) \downarrow -l(a)$, además por la ecuación (1) del capítulo 1, tenemos que $c_{m\,n} < l(a_m) < l(a)$.

En resumen la observación 3 nos dice que $c_{m\,n}$ tiende a $-l(a)$.

De las tres observaciones y del teorema sobre sucesiones dobles podemos concluir lo siguiente,

$$Lim_{m\longrightarrow\infty}\, Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_m)\right) = -l(a)$$

y el que nos interesa:

$$Lim_{n\longrightarrow\infty}\, Lim_{m\longrightarrow\infty}\tfrac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na_m)\right) =$$
$$Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\left(P(X_1 + ... + X_n \geq na)\right) = -l(a).$$

◇

4.1.1. Ejemplos

A continuación ilustramos la teoria de esta sección con ejemplos:

Ejemplo 1 *Si tenemos x una variable aleatoria con la propiedad de que $P(X = c) = 1$ y sea $a > c$ entonces ocurre lo siguiente:*

1. $l(c) = 0$ y 2. $l(a) = \infty$

Las observaciones anteriores se siguen de que
$X\Theta - \ln(E(e^{c\Theta})) = X\Theta - c\Theta = (X - c)\Theta$ y cuando Θ tiende a infinito, la ecuación anterior diverge para $X = a$, mayores que c, y es igual a cero para $X = c$. Obsérvese cómo se satisfacen las cotas dadas en la proposición 2.

Ejemplo 2 *Para la binomial con parámetros (n,p) su función generadora de momentos es $(pe^{\Theta} + 1 - p)^n$, definida en todo $\Re$ con esperanza np*

Para calcular $l(a)$ con $a \in (np, n)$ derivamos a $f(\Theta) = a\Theta - \ln((pe^{\Theta} + 1 - p)^n)$ que es:

$$f'(\Theta) = a - n\frac{pe^{\Theta}}{(pe^{\Theta}+1-p)}$$

y como sólo necesitamos encontrar una raíz, igualamos a cero:

$$a[pe^{\Theta} + 1 - p] - npe^{\Theta} = ape^{\Theta} + a(1 - p) - npe^{\Theta} = 0$$

$$e^{\Theta}(ap - np) + a(1 - p) = 0$$

Nótese cómo la Θ que satisface la ecuación de arriba existe y es única, ya que hacer tender Θ a menos infinito la ecuación de arriba tiende a $a(1 - p) > 0$ y al hacerla tender a infinito se tiene que diverge a menos infinito por
$ap - np = a(n - p) < 0$ y $e^{\Theta}(ap - np) + a(1 - p)$ es estrictamente decreciente.
De hecho tenemos:

$$\Theta^* = \ln\left(\frac{a(1-p)}{p(n-a)}\right)$$

Que, como esperábamos, tenemos a Θ como función continua de a.

Observaciones: Si $a = np$, $\ln\left(\frac{np(1-p)}{p(n-np)}\right) = \ln(1) = 0$ y Si $a \longrightarrow n^-$, la función diverge a menos infinito.

Que es lo mismo que nos confirma la teoría cuando $P(x > n) = 0$.

Finalmente, reduciendo términos:

$$l(a) = aln\left(\frac{a(1-p)}{p(n-a)}\right) - \ln\left(\frac{a(1-p)}{(n-a)} + 1 - p\right)^n$$

Un caso particular es la Bernoulli,toma los valores 0 y 1, $(n, p) = (1, \frac{1}{2})$ con:

$$\Theta^* = \ln(\tfrac{a}{(1-a)})$$

Donde $l(a) = a\ln(a) + (1-a)\ln(1-a) + \ln(2)$ cuando $a \in (1/2, 1]$

que es una función creciente, continua y diferenciable.

Que como se observa: $l(1) = \ln(2)$ y $l(c) = \infty$ cuando $c > 1$

En estos últimos dos casos $l(\cdot)$ se alcanza en el límite de $f(\Theta)$

$$Lim_{\Theta \longrightarrow \infty}\, 2\Theta - \ln\left(\tfrac{1}{2}e^{\Theta} + \tfrac{1}{2}\right)$$

Ejemplo 3 *Para la Poisson con parámetro $\lambda > 0$, su función generadora es $M(\Theta) = e^{\lambda(e^{\Theta}-1)}$ con esperanza λ*

Sea $a \in (\lambda, \infty)$, $f(\Theta) = a\Theta - \ln(e^{\lambda(e^{\Theta}-1)}) = a\Theta - \lambda(e^{\Theta} - 1)$.
Derivando e igualando a cero tenemos:

$$a - \lambda e^{\Theta} = 0$$

Que nos da una única:

$$\Theta^* = \ln(\tfrac{a}{\lambda})$$

que es una función continua en el intervalo (λ, ∞), además de tener inversa continua, $a = \lambda e^{\Theta}$

Por último, $l(a) = a\ln(\frac{a}{\lambda}) - a - 1$ es creciente, continua y diferenciable en el intervalo (λ, ∞)

Los siguientes resultados buscan encontrar condiciones que permitan relacionar convergencia de la función generadora de momentos y la función tasa.

Ejemplo: Si $Lim_{n \longrightarrow \infty} np_n = \lambda$, sabemos que las sucesión de funciones $\{(pe^t + 1 - p)^n\}_{n=1}^{\infty}$ converge a $e^{\lambda(e^{\Theta}-1)}$. Además:

$$Lim_{n \longrightarrow \infty} \ln(\tfrac{a(1-p)}{p(n-a)}) = \ln(\tfrac{a}{\lambda}),$$

$$Lim_{n \longrightarrow \infty} aln(\tfrac{a(1-p)}{p(n-a)}) - \ln(\tfrac{a(1-p)}{(n-a)} + 1 - p)^n = a\ln(\tfrac{a}{\lambda}) - a - 1$$

Nota: El cálculo de los límites anteriores se basa en el siguiente resultado, omitimos su demostración:

Si la sucesión de números $\{c_n\}_{n=1}^{\infty}$ converge a c, entonces

$$Lim_{n\longrightarrow\infty}\left(1-\frac{c_n}{n}\right)^n = e^c$$

c puede tomar el valor ∞, cuando la sucesión es de números reales.

Proposición 12 *Sean $\{X_n\}_{n=1}^{\infty}$, Z variables aleatorias con funciones generadoras de momentos $\{M_n(\Theta)\}_{n=1}^{\infty}$, $M(\Theta)$ respectivamente; todas definidas en D con $0 \in D^o$. Supongamos*

$$Lim_{n\longrightarrow\infty} M_n(\Theta) = M(\Theta) \quad \textit{para toda } \Theta \in D$$

entonces

(i) X_n converge en distribución a z.

(ii) $E(X_n)$ converge a $E(Z)$, de hecho $E(X_n^k)$ converge a $E(Z^k)$, para toda $k \in \mathbb{N}$

(iii) Si $a > E(Z)$ y $P(Z > a) > 0$, entonces

(I) Existe una $m \in \mathbb{N}$ tal que si $n \geq m$ existe Θ_n^*

(II) La sucesión $\{\Theta_n^*\}_{n=m}^{\infty}$ converge a Θ^*, donde $\{\Theta_n^*\}_{n=m}^{\infty}$, Θ^* son donde se maximizan las funciones: $f_n(\Theta) = \{a\Theta - \ln\left(M_n(\Theta)\right)\}_{n=m}^{\infty}$, $f(\Theta) = a\Theta - \ln\left(M(\Theta)\right)$ respectivamente.

Comentario: Recuérdese que las hipótesis $a > E(Z)$ y existe $P(Z > a) > 0$, se puede sustituir por la existencia de Θ^*.

Demostración. (i) Esto se puede consultar en el libro de Patrick Billingsley, Probability and Measure páginas 278 y 390.

(ii) Tomemos $c > 0$ tal que $(-c, c) \subset D$, por el teorema $A.1$ se sigue que $M_n^k(\Theta)$ converge a $M^k(\Theta)$ para toda $\Theta \in (-c, c)$; en particular para $\Theta = 0$ y $k = 1$: $E(X_n^k)$ converge a $E(X^k)$

(iii)

(I) Existencia de las Θ_n^*,

Las Θ_n^* no necesariamente existen para toda $n \in \mathbb{N}$, sin embargo a partir de cierta m está garantizada su existencia por la siguientes dos razones:

1 Por (ii) sabemos: $E(X_n)$ converge a $E(Z)$ y como $E(Z) < a$, existe m_1 tal que si $n \geq m_1$, tenemos $E(X_n) < a$.

2 Como $P(Z > a) > 0$, existe $b > a$ tal que $P(Z > b) > 0$. Sea c que satisfaga que $P(X \in [b, c]) > 0$. Tomemos ξ_1, ξ_2 con las siguientes propiedades $a < \xi_1 < b < c < \xi_2$ y $P(Z = \xi_1) = P(Z = \xi_2) = 0$, que existen por tener los intervalos (a, b), (c, ∞) una cantidad no numerable de reales y sólo un numero numerable de reales, α, con la propiedad $P(Z = \alpha) > 0$.

De la discusión anterior podemos decir,

$P(X_n \leq \xi_1)$ converge a $P(X \leq \xi_1)$,

$P(X_n \leq \xi_2)$ converge a $P(X \leq \xi_2)$ y

$P(Z \leq \xi_1) < P(Z \leq \xi_2)$.

Tomemos m_2 tal que si $n \geq m_2$ se sigue que $P(X_n \leq \xi_1) < P(X_n \leq \xi_2)$. Esto implica que si $n \geq m_2$, $P(X_n > \xi_2) > 0$.

Ahora podemos afirmar:

si $n \geq max\{m_1, m_2\}$, entonces $a > E(X_n)$ y $P(X_n > a) > 0$.

que garantiza la existencia de las Θ_n^*, por la proposición 8.

A partir de ahora supondremos que $E(X_n) = E(Z)$ y $P(X_n > a) > 0$, en otro caso tomamos $m \in \mathbb{N}$ tal que $E(X_n) < a$ y $P(X_n > a) > 0$ para toda $n \geq m$. Sólo se requiere para suponer la existencia del máximo de las funciones $f_n(\cdot)$, que podemos por el inciso anterior.

Demostración de (II) Sea $\varepsilon > 0$, entonces sabemos que existe $\frac{1}{k_\varepsilon} \in \mathbb{N}$ tal que $0 < \frac{1}{k_\varepsilon} < \varepsilon$

Definamos a $\Theta_1 = \Theta^* - \frac{1}{k_\varepsilon}$ y $\Theta_2 = \Theta^* + \frac{1}{k_\varepsilon}$

Por continuidad de la función logaritmo,

para todo $\delta > 0$ existen $n_i := n_i^\delta$ con $i = 1, 2, 3$ tal que si $n \geq n_i$ tenemos

$$\begin{aligned} |f(\Theta_i) - f_n(\Theta_i)| &= |a\Theta_i - \ln(M(\Theta_i)) - [a\Theta_i - \ln(M_n(\Theta_i))]| = \\ &= |\ln(M_n(\Theta_i)) - \ln(M(\Theta_i))| < \delta. \end{aligned}$$

Como lo que queremos saber es que el máximo de las $f_n(\Theta)$ se encuentra en el intervalo $[\Theta^* - k_\varepsilon, \Theta^* + k_\varepsilon]$, definamos a δ como sigue:

$$\delta := max_{i=1,2} \frac{f(\Theta^*) - f(\Theta_i)}{5},$$

y a $m := max_{i=1,2}\{n_i^\delta\}$.

Por concavidad de las f_n y como:

(1) $|f(\Theta^*) - f_n(\Theta^*)| < \delta$

(2) $|f(\Theta_i) - f_n(\Theta_i)| < \delta$ con $i = 1, 2$,

(3) $f(\Theta_i) + 4\delta < f(\Theta^*)$ con $i = 1, 2$,

tomando siempre $n \geq m$,

se sigue $f_n(\Theta_i) < f_n(\Theta^*)$ con $i = 1, 2$.

En otras palabras, el máximo de las $f_n(\Theta)$, Θ_n^*, con $n \geq m$ se encuentra en el intervalo $[\Theta^* - k_\varepsilon, \Theta^* + k_\varepsilon] \subset (\Theta^* - \varepsilon, \Theta^* + \varepsilon)$.

Esto concluye el inciso (II)

$\diamond$

La convergencia de la función tasa es un poco más complicada como lo muestran los siguientes resultados.

Proposición 13 *Sea a tal que* $a > E(Z)$ *y* $P(Z > a) > 0$. *Si* $M_n(\Theta) \longrightarrow M(\Theta)$, *entonces*

$$l(a) \leq Lim\,inf\, l_n(a)$$

Demostración.

Por convergencia puntual tenemos: $f_n(\Theta^*)$ converge a $f(\Theta^*)$ y como en Θ_n^* se alcanza el máximo de la función $f_n(\cdot)$ tenemos $f_n(\Theta^*) \leq f_n(\Theta_n^*)$, tomando límites inferiores de ambos lados:

$$l(a) = f(\Theta^*) \leq Lim\,inf\, f_n(\Theta_n^*) = Lim\,inf\, l_n(a)$$

◇

Observación 10 *La proposición anterior continua siendo válida si substituimos la hipótesis* $P(X > a) > 0$, *por* $P(X = a) > 0$.

Demostración.

Queda claro que basta probar el caso $P(X > a) = 0$ y $P(X = a) > 0$. Para esto recordemos que $Lim_{\Theta \longrightarrow \infty} f(\Theta) = l(a) < \infty$. Sea Θ_ε que satisfaga con $|f(\Theta_\varepsilon) - l(a)| < \frac{\varepsilon}{2}$ y m tal que si $n \geq m$ tenemos $|f_n(\Theta_\varepsilon) - f(\Theta_\varepsilon)| < \frac{\varepsilon}{2}$. Por desigualdad del triángulo tenemos $|f_n(\Theta_\varepsilon) - l(a)| < \varepsilon$, en particular $l(a) - \varepsilon < f_n(\Theta_\varepsilon)$. Tomando el supremo en las funciones $f_n(\cdot)$:

$$l(a) - \varepsilon < l_n(a)$$

Por lo que podemos concluir, para toda $\varepsilon > 0$ existe $m_\varepsilon \in \mathbb{N}$ tal que si $n \geq m$, entonces $l(a) - \varepsilon < l_n(a)$. En otras palabras,

$$l(a) \leq Lim_{n \longrightarrow \infty}\, inf\, l_n(a)$$

◇

La proposición siguiente mejora el resultado, pero requiere de exigir dentro de nuestras hipótesis, convergencia uniforme.

Proposición 14 *Sea a tal que* $a > E(Z)$ *y* $P(Z > a) > 0$. *Si tenemos una de las siguientes dos condiciones:*

(a) La sucesción de funciones $\{M_n(\Theta)\}_{n=1}^{\infty}$ *converge monótonamente a* $M(\Theta)$.

(b) Existe un invtervalo abierto, V, *alrededor de* Θ^* *tal que* $\{M_n(\Theta)\}_{n=1}^{\infty}$ *convergen uniformemente a* $M(\Theta)$ *en este intervalo.*

Entonces

la sucesión $\{l_n(a)\}_{n=1}^{\infty}$ *converge a* $l(a)$. *Las funciones tasa de las variables aleatorias* $\{X_n\}_{n=1}^{\infty}$, Z, *respectivamente.*

(iv) Si además $M_n(\Theta) \downarrow M(\Theta)$ *para toda* $\Theta \in D$, *entonces para a tal que* $P(X = a) > 0$ *y* $P(X > a) = 0$, $l_n(a)$ *converge a* $l(a)$

En el de (a) por el teorema de Dini (véase página 119, del Burkill and Burkill) notamos que $M_n(\Theta)$ converge uniformemente a $M(\Theta)$ en D.

En cualquier caso, sea $K := [\Theta^* - k, \Theta^* + k]$ un intervalo como en la demostración anterior, con $K \subset V$ si ocurre (b).

Por convergencia uniforme tenemos:

(A) Para toda $\varepsilon > 0$ existe m_1 tal que si $n \geq m_1$ tenemos $|f_n(\Theta) - f(\Theta)| < \frac{\varepsilon}{7}$ $\forall \Theta \in K$,

y por la continuidad de $f(\cdot)$:

(B) Para toda $\varepsilon > 0$ existe $\delta > 0$ tal que si $|\Theta - \Theta^*| < \frac{\varepsilon}{7}$,

Sea m_2 tal que si $n \geq m_2$ entonces $|\Theta_n^* - \Theta^*| < \delta$, que existe por (I).

Ahora si $m := max_{i=1,2}\{m_i\}$, tenemos que para toda $n \geq m$

$$\begin{aligned} |f_n(\Theta_n^*) - f(\Theta^*)| &\leq |f_n(\Theta_n^*) - f(\Theta_n^*)| + |f(\Theta_n^*) - f(\Theta^*)| < \\ &< \tfrac{\varepsilon}{7} + |f(\Theta_n^*) - f(\Theta^*)| < \tfrac{2\varepsilon}{7} < \varepsilon \end{aligned}$$

El mayor igual es consecuencia de la desigualdad del triángulo, la primera desigualdad es por (A) y la segunda es por (B).

En resumen,

$\forall \varepsilon > 0$ existe $m \in \mathbb{N}$ tal si $n \geq m$, entonces

$|l_n(a) - l(a)| = |f_n(\Theta_n^*) - f(\Theta^*)| < \varepsilon$.

Demostración de (iv)

A diferencia de (iii) en este caso no existe Θ^*.

De $M_n(\Theta) \downarrow M(\Theta)$ se sigue que
$f_n(\Theta) = a\Theta - \ln(M_n(\Theta)) \uparrow f(\Theta) = a\Theta - \ln(M(\Theta))$, en particular,
$f_n(\Theta) \leq f_{n+1}(\Theta)$.

Tomando supremos, $l_n(a') \leq l_{n+1}(a')$ para toda n natural y para toda $a' \in (E(z), a)$.

Tomando una sucesión $\{a_m\}_{m=1}^{\infty} \subset (E(z), a)$, tenemos que

$$Lim_{m\longrightarrow\infty} Lim_{n\longrightarrow\infty} l_n(a_m) = Lim_{m\longrightarrow\infty} l(a_m) = l(a)$$

y dado que las sucesiones $\{l_n(a_m)\}_{n=1}^{\infty}$, $\{l_n(a_m)\}_{m=1}^{\infty}$ son crecientes, podemos concluir,

$$Lim_{m\longrightarrow\infty}Lim_{n\longrightarrow\infty}l_n(a_m) = Lim_{n\longrightarrow\infty}Lim_{m\longrightarrow\infty}l_n(a_m) = Lim_{n,m\longrightarrow\infty}l_n(a_m) = l(a)$$

En particular,

$$Lim_{n\longrightarrow\infty}Lim_{m\longrightarrow\infty}l_n(a_m) = Lim_{n\longrightarrow\infty}l_n(a) = l(a)$$

$\diamond$

Nota. Al final de la demostración anterior requerimos un resultado de sucesiones dobles, a continuación enunciamos este resultado. Su demostración se puede consultar en el Burkill and Burkill, página 97.

Teorema 2 *Supongamos que cada $x_{n\,m}$ es un real y que para toda m las sucesiones $\{x_{n\,m}\}_{n=1}^{\infty}$ es creciente (decreciente) y para toda n las sucesiones $\{x_{n\,m}\}_{m=1}^{\infty}$ es creciente (decreciente). Si uno de los siguientes limites existe*

$$Lim_{n\,m\longrightarrow\infty}x_{n\,m} \quad , Lim_{n\longrightarrow\infty}x_{n\,m} \quad , Lim_{m\longrightarrow\infty}x_{n\,m}$$

entonces los otros dos también existen y todos ellos son iguales.

Observación 11 *Sean $\{X_n\}_{n=1}^{\infty}$: $P(X_n = 0) = \frac{1}{2}$ y $P(X_n = a_n) = \frac{1}{2}$ donde la sucesión $\{a_n\}_{n=1}^{\infty}$ es monótona creciente y converge a 1 y z la Bernoulli.*

Sus funciones generadoras correspondientes son $\{M_n(\Theta) = \frac{1}{2} + \frac{1}{2}e^{a_n\Theta}\}_{n=1}^{\infty}$, $M(\Theta) = \frac{1}{2} + \frac{1}{2}e^{a_n\Theta}$. De lo que se observa que $M_n(\Theta) \uparrow M(\Theta)$ en los reales positivos; sin embargo con $a = 1$: $l_n(1) = \infty$, si $a_n < 1$, $l(1) = \ln(2)$.

$$l(1) < Lim_{n\longrightarrow\infty}l_n(1)$$

Aún con convergencia mónotona de las funciones generadoras de momentos no podemos garantizar convergencia de las funciones tasa en los puntos donde $P(X = a) > 0$ y $P(X > a) = 0$.

Ejemplo 4 *Para la uniforme $(0, b)$, su función generadora está dada por $\frac{e^{b\Theta}-1}{b\Theta}$ con esperanza $M(\Theta) = \frac{b}{2}$*

Sea $a \in (\frac{b}{2}, b)$

Tenemos que $f(\Theta) = a\Theta - \ln(\frac{e^{b\Theta}-1}{b\Theta}) = a\Theta + \ln(b\Theta) - \ln(e^{b\Theta} - 1)$, derivando nos da:

$$f' = \frac{e^{b\Theta}b\Theta(a-b)+be^{b\Theta}+ab\Theta-b}{b\Theta(e^{b\Theta}-1)}$$

Que si bien no es fácil calcular una raíz, tenemos por observación y proposición 1, además de la proposición 4 garantizada la existencia de esta raíz, más aún por la proposición 8, podemos decir que para este caso $l(a)$ es una función continua en a; pero no lo es si la definimos en todo $\Re$.

Ejemplo 5 *Para la distribución gama con parámetros* (s, λ) $\lambda > 0$*, su función generadora de momentos es* $M(\Theta) = (\frac{\lambda}{\lambda - \Theta})^s$*, definida en* $(-\infty, \lambda)$ *y la esperanza está dada por* $\frac{s}{\lambda}$.

Se tiene $f(\Theta) = a\Theta - s\ln(\frac{\lambda}{\lambda-\Theta}))$, derivando e igualando a cero tenemos $0 = a - \frac{s}{\lambda - \Theta}$ y despejando $\Theta^* = \lambda - \frac{s}{a}$ esto nos dice que

$$l(a) = a(\lambda - \tfrac{s}{a}) - \ln(\tfrac{a\lambda}{s})$$

.

4.1.2. Teorema

Con la proposición 11 dimos una pequeña extensión al teorema del capítulo 1 debilitando la existencia de un óptimo, en esta sección extenderemos este mismo teorema pero ahora será sobre los conjuntos.

Teorema 6 *Consideremos una sucesión* $\{X_i\}_{i=1}^{\infty}$ *de variables aleatorias independientes e idénticamente distribuidas con* $M(\Theta) < \infty$*, entonces*

$$Lim_{n \longrightarrow \infty} \frac{1}{n} \ln\Big(P(\frac{X_1 + ... + X_n}{n} \in A)\Big) = -inf_{z \in A} l(z) \tag{4.13}$$

Donde A *es de la forma*

i) $[a, \infty)$ *si* $a > E(x)$.

ii) (a, ∞) *si* $a > E(x)$.

iii) $(a, b), [a, b], [a, b), (a, b]$ *si* $a < b$.

En el caso que $inf_{z \in A} l(z) \neq \infty$*, tenemos* $inf_{z \in A} l(z) = l(a)$.

Demostración.

El caso (i) es el teorema del capítulo 1.

Para el caso (ii), notamos que:

Si $P(X > a) > 0$, entonces $l(\cdot)$ es continua en a por proposición 8, además,

$$\frac{1}{n} \ln\Big(P(\frac{X_1 + ... + X_n}{n} \geq a)\Big) \leq \frac{1}{n} \ln\Big(P(\frac{X_1 + ... + X_n}{n} \in (a, \infty))\Big) \tag{4.14}$$

$$\frac{1}{n} \ln\Big(P(\frac{X_1 + ... + X_n}{n} \in (a, \infty))\Big) \leq \frac{1}{n} \ln\Big(P(\frac{X_1 + ... + X_n}{n} \in [a_m, \infty))\Big) \text{ donde } a_m \downarrow a \tag{4.15}$$

Ahora como $l(\cdot)$ es creciente y por 4.14 y 4.15 tenemos,

$$-l(a) \leq Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\Big(P(\tfrac{X_1+...+X_n}{n}\in(a,\infty))\Big)$$

$$Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\Big(P(\tfrac{X_1+...+X_n}{n}\in(a,\infty))\Big) \leq -l(a_m)\ \forall m\in\mathbb{N}$$

aplicando limite sobre m,

$$Lim_{n\longrightarrow\infty}\tfrac{1}{n}\ln\Big(P(\tfrac{X_1+...+X_n}{n}\in(a,\infty))\Big) = -l(a)$$

en otro caso, es decir $P(X>a)=0$, todo se reduce a tener $-\infty$ de ambos lados de la ecuación (4.13), por conclusión 1; en cualquier caso :

$$Lim_{n\longrightarrow\infty}\frac{1}{n}\ln\Big(P(\frac{X_1+...+X_n}{n}\in(a,\infty))\Big) = -l(a) \tag{4.16}$$

De nuevo por ser $l(\cdot)$ creciente $l(a)=inf_{z\in(a,\infty)}l(z)$.

Para el caso (iii), Si $P(X>a)=0$ es trivial, en otro caso tenemos lo siguiente

$$P\Big(\frac{X_1+...+X_n}{n}\in(a,b)\Big) = P\Big(\frac{X_1+...+X_n}{n}\in(a,\infty)\backslash[b,\infty)\Big) = P\Big(\frac{X_1+...+X_n}{n}\in(a,\infty)\Big) - P\Big(\frac{X_1+...+X_n}{n}\in[b,\infty)\Big) \tag{4.17}$$

Una observación que requerimos sobre límites es la siguiente:

Sea $Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(c_n)=c$, $Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(b_n)=b$ y $b\leq c$

Si $Lim_{n\longrightarrow\infty}\sqrt[n]{1-\frac{b_n}{c_n}}=1$, entonces $Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(c_n-b_n)=c$.

Por consiguiente usando la observación anterior y (4.17),

$Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(P(\frac{X_1+...+X_n}{n}\in(a,b))) = Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(P(\frac{X_1+...+X_n}{n}\in (a,\infty)\backslash[b,\infty))) =$
$Lim_{n\longrightarrow\infty}\frac{1}{n}\ln[P(\frac{X_1+...+X_n}{n}\in(a,\infty)) - P(\frac{X_1+...+X_n}{n}\in[b,\infty))] =$

$= Lim_{n\longrightarrow\infty}\frac{1}{n}\ln[P(\frac{X_1+...+X_n}{n}\in(a,\infty))] = -l(a)$

La última igualdad se debe a la observación de limites:
$P(\frac{X_1+...+X_n}{n}\in(a,\infty)) = e^{-l(a)+o(n)}$, $P(\frac{X_1+...+X_n}{n}\in[b,\infty)) = e^{-l(b)+o(n)}$ con $l(a)<l(b)$,

$$Lim_{n\longrightarrow\infty}\sqrt[n]{1-\frac{P(\frac{X_1+...+X_n}{n}\in[b,\infty))}{P(\frac{X_1+...+X_n}{n}\in(a,\infty))}} = 1 \tag{4.18}$$

Nótese que por ser $l(\cdot)$ creciente y continua, $l(a) = inf_{z\in(a,b)} l(z)$; el otro caso es trivial

Debe observarse que como $P(X > a) > 0$ existe una m tal que si $n \geq m$ tenemos que $P(\frac{X_1+...+X_n}{n} \in (a,b)) > 0$, para que el logaritmo tenga sentido. Los demás casos se resuelven de manera análoga.

$\diamond$

En el caso de tomar el ejemplo 2 con la Bernoulli y $\{a = \frac{2}{3}\}$ tenemos que $Lim_{4n\longrightarrow\infty} \frac{1}{4n} \ln(P(\frac{X_1+...+X_{4n}}{4n} \in \{\frac{2}{3}\})) = -\infty$, pero $0 < inf_{a\in\frac{2}{3}} l(\frac{2}{3}) < \infty$

Proposición 15 *Sea* $Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n) = c$, $Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(b_n) = b$ *y* $b \leq c$
Si $Lim_{n\longrightarrow\infty} \sqrt[n]{1 - \frac{b_n}{c_n}} = 1$, *entonces* $Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n - b_n) = c$.

Demostración.
$Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n - b_n) = Lim_{n\longrightarrow\infty} \ln((c_n - b_n)^{\frac{1}{n}})$
Por continuidad de la función exponencial, podemos sustituir el limite anterior, por el siguiente $Lim_{n\longrightarrow\infty}(c_n - b_n)^{\frac{1}{n}} = Lim_{n\longrightarrow\infty} \sqrt[n]{c_n}\sqrt[n]{1 - \frac{b_n}{c_n}} =$
$Lim_{n\longrightarrow\infty} e^{\frac{1}{n}\ln c_n}\sqrt[n]{1 - \frac{b_n}{c_n}} = e^c$.
Por lo que, $Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n - b_n) = Lim_{n\longrightarrow\infty} \ln((c_n - b_n)^{\frac{1}{n}}) = c$
$\diamond$

Nótese que pedir el limite $Lim_{n\longrightarrow\infty} \sqrt[n]{1 - \frac{b_n}{c_n}} = 1$, no implica que $Lim_{n\longrightarrow\infty} \frac{b_n}{c_n} \neq 1$, como lo muestra el caso siguiente;
$b_n = e^{-n-\frac{1}{n}}$, $c_n = e^{-n}$, $c_n \sim b_n$,
$Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n) = Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(b_n) = -1$; pero
$Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(c_n - b_n) = -1$, su cálculo es seguir los pasos de la demostración anterior.

Es más que pedir que no sean iguales, en el sentido $\sim$, es necesario que no converjan a lo mismo demasiado rápido.

Para el caso de la uniforme en $(0,1)$ resulta claro que para cualquier $a \in (0,1)$

$$Lim_{n\longrightarrow\infty} \frac{1}{n}\ln(P(\frac{X_1+...+X_n}{n} \in \{a\})) = -\infty$$

pero $l(a) \in (0,\infty)$ véase el ejemplo 4.

El siguiente ejemplo ilustra la idea de que aunque para los conjuntos de la forma $\{a\}$, $a \in \Re$ no sea válido el teorema 6, en el caso en el que

$u := P(X = a) > 0$ sí tendremos, $Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(P(\frac{X_1+...+X_n}{n} \in a)) = -l(a)$; esto será por la razón de que aunque
$P(X_1 + ... + X_n > na) \sim P(X_1 + ... + X_n \geq na) \sim e^{-l(a)}$ el factor $o(n)$ de $e^{-l(a)+o(n)}$ pesará lo suficiente para que,

$$Lim_{n\longrightarrow\infty} \sqrt[n]{1 - \frac{P(X_1+...+X_n>na)}{P(X_1+...+X_n\geq na)}} = 1$$

Proposición 16 *Consideremos una sucesión $\{X_i\}_{i=1}^{\infty}$ de variables aleatorias independientes e idénticamente distribuidas, Sea $a > E(X)$ tal que $c := P(X = a) > 0$, entonces*

$$Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(P(\frac{X_1+...+X_n}{n} \in a)) = -l(a)$$

Demostración

La demostración sigue el camino del teorema 6 inciso (iii), aquí solo mostraremos que

$$Lim_{n\longrightarrow\infty} \sqrt[n]{1 - \frac{P(\frac{X_1+...+X_n}{n} \in (a,\infty))}{P(\frac{X_1+...+X_n}{n} \in [a,\infty))}} = 1 \tag{4.19}$$

Nótese que

$P(\frac{X_1+...+X_n}{n} \in [a,\infty)) = P(\frac{X_1+...+X_n}{n} \in (a,\infty)) + P(\frac{X_1+...+X_n}{n} \in \{a\}) \geq$

$P(\frac{X_1+...+X_n}{n} \in (a,\infty)) + c^n$

Definamos $u_n := P(\frac{X_1+...+X_n}{n} \in (a,\infty))$, entonces la ecuación (4.19) se puede acotar por

$$Lim_{n\longrightarrow\infty} \sqrt[n]{1 - \frac{u_n}{u_n + c^n}} = Lim_{n\longrightarrow\infty} \sqrt[n]{\frac{u_n + c^n - u_n}{u_n + c^n}}$$
$$Lim_{n\longrightarrow\infty} \sqrt[n]{\frac{c^n}{u_n + c^n}} = Lim_{n\longrightarrow\infty} \frac{c}{\sqrt[n]{u_n + c^n}}$$

pero u_n tiende a cero conforme n tiende a infinito, por lo que el limite anterior es igual a 1

◇

4.1.3. Aplicación

Esta sección puede ser complementada con la sección de Leyes de los grandes números de los preliminares, aunque no es necesario para su comprensión.

La siguiente aplicación, mostrará como podemos usar el teorema 6 con $A = (a, \infty)$ para demostrar el teorema de los grandes números con la teoría de esta sección. Antes daremos algunos prerrequisitos de teoría de la medida que solo se requieren en esta sección.

Los teoremas 7 y 8 permitiran que problemas sobre convergencia casi-seguramente, puedan responderse conociendo convergencia de sucesiones de números reales no negativos.

Teorema 7 *La sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias, converge a X P-c.s si y sólamente si*

$$Lim_{n \longrightarrow \infty} P(sup_{n \geq m}|X_n - X| > \varepsilon) = 0 \text{ para toda } \varepsilon > 0$$

Demostración.

La sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias converge a X P-c.s si y sólamente si el conjunto A^c definido como

$$A^c = \cup_{k \geq 1} \cap_{m \geq 1} \cup_{n \geq m} \{|X_n - X| > \tfrac{1}{k}\} = \cup_{k \geq 1} \cap_{m \geq 1} \{sup_{n \geq m}|X_n - X| > \tfrac{1}{k}\}$$

satisface que $P(A^c) = 0$, es decir

$$P(\cap_{m \geq 1}\{sup_{n \geq m}|X_n - X| > \tfrac{1}{k}\}) = Lim_{m \longrightarrow \infty} P(\{sup_{n \geq m}|X_n - X| > \tfrac{1}{k}\}) = 0$$

y esto para cualquier $k \in \mathbb{N}$, en particular para toda ε podemos tomar $k := k_{\varepsilon}$ tal que $\varepsilon > \frac{1}{k}$

$\diamond$

Definición 1 *La sucesión $\{X_n\}_{n=1}^{\infty}$ converge P-estocásticamente a X, cuando*

$$\textstyle\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty \text{ para toda } \varepsilon > 0.$$

Teorema 8 *Si $\{X_n\}_{n=1}^{\infty}$ converge x P-estocásticamente, entonces $\{X_n\}_{n=1}^{\infty}$ converge a x P-c.s.*

Demostración.

Por el Lema de Borel-Cantelli, se sigue,

$$P(Lim\, sup\{|X_n - X > \varepsilon|\}) = P(|X_n - x > \varepsilon| \text{ una infinidad de veces})$$

esto es equivalente,

$Lim_{m \longrightarrow \infty} P(sup_{n \geq m}|X_n - X| > \varepsilon) = 0$

por la siguiente razón:

Sean $B_m := \cup_{n=m}^{\infty}\{|X_n - X| > \varepsilon\}$ y $B := Lim\,sup\{|X_n - X| > \varepsilon\}$, por construcción $B_n \downarrow B$, $Lim_{m\longrightarrow\infty}P(B_m) = P(B)$ y
$B_m = sup_{n\geq m}\{|X_n - X| > \varepsilon\}$.

Y usando el teorema 7 tenemos convergencia casi seguramente

Los teoremas 7 y 6, con $A = (\varepsilon, \infty)$, nos dan como consecuencia la Ley fuerte de los grandes números.
Sin pérdida de generalidad $E(X) = 0$.

$$P(|\frac{X_1 + ... + X_n}{n}| > \varepsilon) = P(\frac{X_1 + ... + X_n}{n} > \varepsilon) + P(\frac{X_1 + ... + X_n}{n} < -\varepsilon) = \\ = P(\frac{X_1 + ... + X_n}{n} \in (\varepsilon, \infty)) + P(\frac{X_1 + ... + X_n}{n} \in (-\infty, -\varepsilon)) \tag{4.20}$$

Por teorema 6 (ii),

$$(4.20) \sim e^{-nl(\varepsilon)+o(n)} + e^{-nl(-\varepsilon)+o(n)}$$

o en su caso,

$$(4.20) \leq e^{-nl(\varepsilon)+o(n)} + e^{-nl(-\varepsilon)+o(n)}$$

cuando $l(\varepsilon) = \infty$ o $l(-\varepsilon) = \infty$.
De cualquier manera,

$$\sum_{n=1}^{\infty} P(|\frac{X_1 + ... + X_n}{n}| > \varepsilon)$$

converge si,

$$\sum_{n=1}^{\infty} e^{-nl(\varepsilon)+o(n)} + e^{-nl(-\varepsilon)+o(n)}$$

lo hace.
Pero esta es una sucesión convergente, por criterios de comparación en serier, obsérvese: $n^2e^{-nc+o(n)}$ converge a cero y la serie $\sum_{n=1}^{\infty}\frac{1}{n^2} < \infty$. Y se satisface el teorema 8, sobre convergencia P-estocásticamente. Que conluye, $|\frac{X_1+...+X_n}{n}|$ converge a 0 casi seguramente.
$\diamond$
Algunas observaciones son: el tomarse rayos abiertos requiere que $l(\cdot)$ sea continua, que a su vez requiere que $M(\Theta)$ sea finita en algún intervalo alrededor del cero, en particular existen todos los momentos, $E(X^n) < \infty$ para todo $n \in \mathbb{N}$. El modificar la demostración con intervalos cerrados no nos permite liberanos de estas condiciones por la razón de que tanto en la cota superior como en la inferior del teorema del capítulo 1 hacemos uso de que $M(\Theta) < \infty$ en un intervalo alrededor del cero.
Nuestra ley fuerte de los grandes números quedaría,

Teorema 9 *Sea* $\{X_n\}_n^\infty$ *una sucesión de variables aleatorias independientes e idénticamente distribuidas, tal que* $M(\Theta)$ *este definida en un intervalo alrededor del cero; entonces*

$$\left(\frac{X_1 + \ldots X_n}{n}\right)_{n\in\mathbb{N}} \quad \textit{converge a } E(X) \; P\textit{-casi-seguramente.}$$

Capítulo 5

Grandes Desviaciones para Proceso de Ramificación

El presente capítulo relaciona las ideas de los capítulos anteriores con el proceso de Ramificación. El capítulo puede dividirse en tres bloques, el primero trata distintas variaciones del teorema principal del capítulo 1, la segunda parte trata el modelo de Galton-Watson y por último relacionamos las dos partes anteriores.

Teorema 10 *Si $a > E(X)$, entonces*

$$P(X_1 + ... + X_n \geq ja \ \forall j \in \{1, ..., n\}) = e^{-nl(a)+o(n)} \tag{5.1}$$

si y sólo si

$$P(X_1 + ... + X_n \geq na) = e^{-nl(a)+o(n)}. \tag{5.2}$$

La demostración requiere del siguiente lema.

Lema 1 *Supongamos que los números reales $r_1, ..., r_n$ satisfacen $r_1 + ... + r_n \geq n\beta$ y sea $i^* = inf\{x : x = r_1 + ... + r_i - \beta \text{ con } i \in \{1, ..., n\}\}$, entonces*

$$r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} \geq j\beta \ \text{ para todo } \ 1 \leq j \leq n.$$

Debe entenderse a $(i)_n = i \ mod \ n$.

Demostración del lema.

Definamos a k como $i^* + k = n$

Si $k = 0$, entonces $i^* = n$ y $r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} = r_{(1)} + ... + r_{(j)}$.

Supongamos que $r_{(1)} + ... + r_{(j)} - j\beta < 0$ para alguna $j \in 1, ..., n$. De lo que se sigue $r_{(1)} + ... + r_{(j)} - j\beta < 0 \leq r_{(1)} + ... + r_{(n)} - n\beta$ que contradice que $i^* = n$

Si $k > 0$.

Aquí será por casos, si $j \in \{1, ..., k\}$, es decir $i^* < i^* + j \leq n$.

También tenemos $(i^* + s)_n = i^* + s$ si $s \leq j$.

Si $r_{(i^*+1)} + ... + r_{(i^*+j)} - j\beta < 0$, entonces:

$$r_{(1)} + ... + r_{(i^*+j)} - (i^* + j)\beta =$$
$$\big(r_{(1)}+...+r_{(i^*+j)}-(i^*)\beta\big)+\big(r_{(i^*+1)}+...+r_{(i^*+j)}-j\beta\big) < r_{(1)}+...+r_{(i^*+j)}-(i^*)\beta$$

Es decir, $r_{(1)} + ... + r_{(i^*+j)} - (i^* + j)\beta < r_{(1)} + ... + r_{(i^*+j)} - (i^*)\beta$
que es una contradicción por definición de i^*.

Ahora, sea $j > n$.

$r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} = r_{(i^*+1)} + ... + r_{(i^*+k)} + r_1 + ... + r_{(j-k)}$.

$$\text{Si, } r_{(i^*+1)} + ... + r_{(i^*+k)} + r_1 + ... + r_{(j-k)} < j\beta, \text{ entonces}$$
$$r_1 + ... + r_{(j-k)} < -a + (j - k)\beta, \text{ donde } a = r_{(i^*+1)} + ... + r_{(i^*+k)}.$$

De lo que se sigue:

$r_1 + ...r_{i^*} + r_{i^*+1} + ...r_{i^*+k} \geq k\beta + (n - k)\beta = k\beta + i^*\beta$, entonces
$r_1 + ...r_{i^*} \geq i^*\beta - a$ Por lo tanto:

$$r_1 + ...r_{i^*} \geq i^*\beta - a$$
$$r_1 + ... + r_{(j-k)} < (j - k)\beta - a$$
$$r_1 + ... + r_{(j-k)} - (j - k)\beta < a \leq r_1 + ...r_{i^*} - i^*\beta$$

Que es una contradicción.

$\diamond$

Demostración del teorema.

Por monotonía de P se tiene:

$$P(X_1 + ...X_n \geq ja \quad \forall j \in \{1, ..., n\}) \leq P(X_1 + ... + X_n \geq na)$$

Por el lema anterior, tenemos que si $x_1 + ... + x_n \geq na$, existe una i^* con

$$r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} \geq ja \text{ para todo } 1 \leq j \leq n$$

de lo que tenemos.

$$P(x_1 + ... + x_n \geq na) = P(r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} \quad \forall j \in \{1, ..., n\}) =$$
$$\sum_{k=1}^{n} P(r_{(i^*+1)_n} + ... + r_{(i^*+j)_n} \geq ja \quad \forall j \in \{1, ..., n\}, k = i^*) \qquad (5.3)$$

Esta última igualdad es consecuencia de que para $k \neq i^*$, los sumandos son cero. Por monotonía:

$$(5.3) \leq n\,P(r_{(k+1)_n} + ... + r_{(k+j)_n} \geq ja \quad \forall j \in \{1,...,n\}) =$$
$$= nP(X_1 + ...X_n \geq ja \quad \forall j \in \{1,...,n\})$$

La última igualdad se debe a que las $\{X_k\}$ son independientes e idénticamente distribuidas.

$\diamond$

Ahora introduciremos una posición inicial X_0 que será independiente de las $\{X_k\}_{k=1}^{\infty}$. Notación, $M_0(\Theta) := E(e^{X_0\Theta})$.

Teorema 11 *Sea $a > E(X_1)$ y $P(X_1 > a) > 0$.*
Si M_0 es finita en una bola abierta que contiene al cero y en todo los reales positivos, entonces

$$P(X_0 + X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) \leq e^{-nl(a)+o(n)} \tag{5.4}$$

Si $P(X \geq 0) > 0$, entonces

$$P(X_0 + X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) \geq e^{-nl(a)+o(n)} \tag{5.5}$$

Demostración. Para $\Theta \geq 0$

$$P(X_0 + X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) \leq$$
$$\leq P(X_0 + X_1 + ... + X_n \geq na) = P(\Theta(x_0 + x_1 + ... + x_n) \geq \Theta(na)) =$$
$$P(e^{\Theta(X_0+X_1+...+X_n)} \geq e^{\Theta(na)}) \leq e^{-\Theta(na)}E(e^{X_0+X_1+...+X_n}) =$$
$$por\ Tchebychev$$
$$= e^{-\Theta(na)}E(e^{X_0})E(e^{X_1})^n = e^{-\Theta(na)}M_0(\Theta)M^n(\Theta)$$
$$e^{-\Theta(na)+\ln(M_0(\Theta)+n\ln(M(\Theta)} = e^{-n(a\Theta-\ln(M(\Theta))+\ln(M_0(\Theta))}$$
$$Como\ P(X_1 > a) > 0\ podemos\ tomar\ \Theta := \Theta^*.$$
$$e^{-nl(a)+\ln(M_0(\Theta^*))}$$

Ahora, si $P(X_0 < 0) = 1$, la parte derecha del teorema 11 es siempre igual a cero, por lo que la cota inferior no se puede dar. En otro caso,

$$P(X_0 + X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) \geq$$
$$\geq P(X_0 \geq 0;\ X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) =$$
$$P(X_0 \geq 0)P(X_1 + ... + X_n \geq ja \quad \forall j \in \{1,...,n\}) =$$
$$= e^{-nl(a)+o(n)}$$

$\diamond$

Corolario 1 *Supongamos que $a > E(X)$, $P(X > a) > 0$ y que M_0 es finita en una bola abierta que contiene al cero y en todo los reales positivos, entonces*

$$P(X_0 + X_1 + ... + X_j > ja \quad \forall j \in 1, ..., n) \leq e^{-nl(a)+o(n)}$$

$$Si\ P(x_0 > 0) > 0,\ entonces$$

$$P(X_0 + X_1 + ... + X_j > ja \quad \forall j \in 1, ..., n) \geq e^{-nl(a)+o(n)}$$

La demostración sigue las líneas del teorema anterior y requiere de la generalización del teorema del capítulo 1 a finales del capítulo anterior. Si en el teorema 11 ponemos a $P(X_1 = 0) = 1$ y $a > 0$, entonces tenemos $l(a) = \infty$, $P(X_0 + X_1 + ... + X_j \geq ja\ \forall j \in \{1, ...n\}) = P(X_0 \geq na)$ y con X_0 una Poisson queda claro que no se puede dar la ecuación (5.4). Para que se satisfaga el teorema 11 al menos es necesario pedir que $P(X \geq a) > 0$.

5.1. Galton- Watson

Consideremos una población de individuos que al final de su vida se divide en nuevos individuos. Por individuos puede entenderse átomos en una fisión nuclear, bacterias en una colonia, personas o errores como se verá más adelante. A continuación presentaremos el proceso de ramificación conocido como Galton-Watson, un modelo sencillo que es usado para entender la dinámica de poblaciones. De particular importancia es la pregunta por la probabilidad de que la población se extinga; esta fue también el origen de la motivación de Francis Galton y el reverendo Watson a finales del siglo 19. En ese entonces se preguntaban sobre la probabilidad de extinción de descendientes masculinos, porque de esto se seguía la desaparición de los títulos nobiliarios, en ese entonces sólo se heredaban los títulos nobiliarios a los hijos varones.

Modelo de Galton-Watson (Proceso de ramificación)

Comenzamos con un único "pariente", nodo; a cada nodo le asignamos de manera aleatoria un número de "hijos", este número seleccionado independientemente de una distribución discreta $\{p_i,\ i = 1, ...\}$. Formalmente, sea $\{y(i, j)\}$ variables aleatorias independientes e idénticamente distribuidas con distribución $p = (\{p_k\})$, representando el número de hijos del nodo j en la generación i. El número Y_n indica el número de parientes en la generación n, y está dada por la ecuación:

$$Y_0 = 1$$

$$Y_n = \sum_{i=1}^{Y_{n-1}} y(n-1, i) \tag{5.6}$$

Denotemos con b la esperanza del número de descendientes, que supondremos finito, es decir, $b = \sum_{i=1}^{\infty} ip_i$. El modelo de Galton-Watson está dado por la cadena de Markov (Y_n) con matriz de transición

$$q_{k,j} = \sum_{j_1+\ldots j_k=j} p_{j_1} * \ldots * p_{j_k} \; k \geq 1, \; j \in \mathbb{N}.$$

Paso de k individuos en una generción a tener j en la siguiente:
Para el caso $k = 0$, en el que la población está extinta tendremos $q_{o,j} = \delta_{0,j}$.
Denotaremos al primer tiempo de extinción,

$$T_0 = inf\{n \geq 1 : \; Y_n = 0\}.$$

Y la probabilidad de extinción cuando la población inicia con i individuos por: $\rho_i = P_i(T_0 < \infty) = P(T_0|Y_0 = i)$. Para eliminar casos triviales en lo siguiente tendremos $i > 0$, $p_o \in (0,1)$ y $p_0 + p_1 < 1$. Denotemos $\rho := \rho_1$, entonces se sigue que

$$\rho_i = \rho^i$$

dado que cada ramificación de un individuo de la primera generación se desarrolla independiente y con la misma distribución que los otros. Lo que nos interesará investigar es el valor de ρ, es decir, nuestra población inicial será un sólo individuo $Y_0 = 1$. Este individuo puede extinguirse (con probabilidad p_0), si no tiene ningún descendiente, o puede tener k descendientes (con probabilidad $p_k = P(Y_1 = k|Y_0 = 1)$). La población $Y_1 = k$ se extingue con la misma probabilidad con la que se extinguiría la población $Y_0 = k$ (p_k), por propiedades de cadenas de Markov. De lo anterior obtenemos la probabilidad de extinción

$$\rho = p_0 + \sum_{k=1}^{\infty} P(Y_1 = k|Y_0 = 1)\rho_k = \sum_{k=1}^{\infty} p_k \rho^k$$

Definamos la función

$$f(s) = \sum_{k=1}^{\infty} p_k s^k \quad s \in [0,1] \tag{5.7}$$

que se observa tiene una propiedad interesante: La probabilidad de extinción ρ es un punto fijo $f(\rho) = \rho$. Obviamente tenemos como punto fijo a 1 pues $f(1) = 1$.

Estimación de la probabilidad de extinción

Para estimar ρ distinguiremos dos casos $f'(1) = \sum_{k=1}^{\infty} k\, p_k = E(p) \leq 1$ y $f'(1) = E(p) > 1$. En ambos casos

$$f''(s) = \sum_{k=2}^{\infty} k(k-1)p_k s^{k-2} > 0 \;\; s \in [0,1]$$

f es estrictamente convexa en el intervalo $[0,1]$, $0 < f(0) < 1$, de lo que tenemos que en el caso de $E(p) \leq 1$ el 1 es el único punto fijo de f, que nos permite decir que la probabilidad de que la población se extinga en este caso es 1. En contraste con el caso $E(p) \leq 1$ esto es cuando $E(p) > 1$, se sigue de la convexidad estricta que f tiene otro único punto fijo, $\tilde{\rho}$, en el intervalo $\tilde{\rho} \in [0,1)$. Lo que mostramos a continuación es que $\rho = \tilde{\rho}$. Para ello sea

$$g_n : P(Y_n = 0|Y_0 = 1), \; n \geq 1$$

la probabilidad de que en la n-ésima generación el proceso esté extinto. Ahora por ser $f^{'} > 0$, la función es estrictamente creciente, por consiguiente

$$g_1 = p_0 = f(0) \leq f(\tilde{\rho}) = \tilde{\rho}$$

del mismo modo tenemos:

$$g_n \leq \tilde{\rho} \text{ se sigue } g_{n+1} = f(g_n) \leq M(\tilde{\rho}) = \tilde{\rho}$$

Por inducción concluimos que $g_n \leq \tilde{\rho}$ para toda $n \in \mathbb{N}$, y por popiedades de cadenas de Markov tenemos $\rho = Lim_{n \longrightarrow \infty} g_n \leq \tilde{\rho}$. Pero como en el intervalo $(0, \tilde{\rho})$ no existe otro punto fijo, se sigue que $\rho = \tilde{\rho}$.

Resumiendo, la esperanza de los hijos que tiene un individuo, $E(p)$, importa sobre la supervivencia de la población. En el caso que $E(p) \leq 1$, la población se terminará extinguiendo. Ocurre el caso que $E(p) > 1$, entonces existirá una probabilidad positiva ρ, pero no igual a 1, de que la población desaparezca y es la raíz más chica de la función f.

5.2. Procesos de ramificación y grandes desviaciones

Teorema 12 *(i)* $E(Y_n) = b^n$

(ii) Si $b \leq 1$, *entonces* $P(Lim_{n \longrightarrow \infty} Y_n = 0) = 1$

(iii) Si $b > 1$, *entonces* $P(Lim_{n \longrightarrow \infty} Y_n = \infty) > 0$

Demostración.

Los últimos dos incisos ya están demostrados. Para (i) usaremos inducción.

Para $n = 1$ es por hipótesis. supongamos que es válido para $n - 1$,

$$E(Y_n) = \sum_{j=0}^{\infty} E(Y_n | Y_{n-1} = j) P(Y_{n-1} = j) = (*)$$

dado que tengo j individuos en la generación $n - 1$ y cada uno tiene cierto número de hijos independientemente de los otros, se sigue

$$(*) = \sum_{j=0}^{\infty} jbP(Y_{n-1} = j) = b \sum_{j=0}^{\infty} jP(Y_{n-1} = j) = bE(Y_{n-1}) = bb^{n-1} = b^n$$

$\diamond$

Extendamos un poco el problema de la siguiente manera. El nodo inicial tiene una propiedad que es hereditaria, puede ser un mal genético, un dote particular y queremos saber si en algún momento desaparecerá esa propiedad, esta propiedad no necesariamente necesita ser binaria, por ejemplo la propiedad de despertarse a cierta hora del día o comer cierta cantidad de un producto, etc. Generalizamos el modelo de Galton-Watson como sigue, comencemos con el proceso de ramificación y asignemos el valor $x(n,i)$ al miembro de i de la generación n, como sigue. El nodo inicial del proceso de ramificación le asignamos el valor x_0. Cada hijo de un individuo se le asigna su valor remplazando él de su ancestro por una variable aleatoria, con distribución F, que es independiente de todas las demás. Formalmente, sea $\{w(n,i)\}$ variables aleatorias idénticamente distribuidas, independientes y con distribución F. En otras palabras, si el miembro i de la generación n es hijo del miembro j de la generación $n-1$, entonces

$$x(n,i) = x(n-1,j) + w(n,i) = x_0 + \sum_{l=1}^{n} w(l, L(l,n,i))$$

donde $L(l,n,i)$ denota al ancestro en la generación l del miembro (n,i). Obsérvese que la construcción nos dice que $\{x(n,1), x(n,2), ...\}$ son idénticamente distribuidas pero no independientes. Denotemos con x_n a la variable con la misma distribución que las de la generación n y w_n a la variable aleatoria con distribución F. Por último, denotemos con $Z(n,x)$ al número de miembros de la generación n que tienen valor positivo cuando $x_0 := x$. Para fijar ideas, si la propiedad de la que hablamos anteriormente la tiene un individuo cuando el valor es estrictamente mayor que cero. Cuando nos preguntamos por los individuos de la generación n que aún poseen la propiedad, en los términos anteriores nos preguntamos por los individuos que a lo largo de su árbol genealógico satisfacen:

$$\{x(j, L(j,n,i)) > 0 \quad \forall j \in \{1, ..., n\}\} \tag{5.8}$$

Lema 2 $E(Z(n,x)) = E(Y_n)P(x_n \geq 0)$

Demostración.

Observación, $E(Z(n,x)) = E(\sum_{i=1}^{Y_n} 1_{\{x(n,i)\geq 0\}})$

Ahora condicionando y por linealidad tenemos que,

$E(\sum_{i=1}^{Y_n} 1_{\{x(n,i)\geq 0\}}|Y_n) = \sum_{i=1}^{Y_n} E(1_{\{x(n,i)\geq 0\}}|Y_n) =$

$= Y_n E(1_{\{x(n,1)\geq 0\}}|Y_n) = Y_n E(1_{\{x(n,1)\geq 0\}})$

La última igualdad se tiene por ser idénticamente distribuidas las $x(n,i)$. Por lo tanto;

$E(Z(n,x)|Y_n) = Y_n E(1_{\{x(n,1)\geq 0\}}|Y_n) = Y_n E(1_{\{x(n,1)\geq 0\}})$

Condicionando,

$$\begin{aligned} E(E(Z(n,x)|Y_n)) = E(Y_n E(1_{\{x(n,1)\geq 0\}}|Y_n)) = E(Y_n)E(1_{\{x(n,1)\geq 0\}}) = \\ = E(Y_n)P(x(n,1) \geq 0) \end{aligned}$$

Conclusión,

$$E(Z(n,x)) = E(Y_n)P(x(n,1) \geq 0)$$

$\diamond$

El teorema 12 nos dice que cuando $b \leq 1$ $Z(n,x)$ se va terminar yendo a cero. El resultado interesante se presenta cuando $b > 1$, los individuos cada vez son más, pero $E(w_n) < 0$, con el tiempo la probabilidad de heredar la propiedad cada vez es menor. Un primer resultado en este sentido, es el siguiente.

Teorema 13 (DE BIGGINS)

Consideremos el caso en el que $b > 1$ y $E(w_n) < 0$

(i) Para cualquier x fija,

$$Lim_{n\longrightarrow\infty}\frac{1}{n}\ln(E(Z(n,x))) = \ln(b) - l(0) \tag{5.9}$$

donde $l(0)$ esta en función de la distribución F.

(ii) Si $\ln(b) - l(0) < 0$, entonces $P(Lim_{n\longrightarrow\infty}Z(n,x) = 0) = 1$

(iii) Si $\ln(b) - l(0) > 0$, entonces $P(Lim_{n\longrightarrow\infty}sup\, Z(n,x) = \infty) > 0$

Demostración.

Por el lema 2 tenemos,

$$\ln(E(Z(n,x))) = n\ln(b) + \ln P(x_n \geq 0)$$

y por el teorema 11, $P(x_n) = e^{-l(a)+o(n)}$, dado que $0 > E(w_n)$,

$$\ln(E(Z(n,x))) = n\ln(b) - l(0) + o(n)$$

y tomando límites queda demostrado (i)

Para (ii), notamos que si $\ln(b) - l(0) < 0$, necesarimente por (i), $\sqrt[n]{E(Z(n,x))}$ converge a un número menor que 1, pero esto me indica que la suceción $E(Z(n,x))$ converge a 0 geometricamente. Ahora, por la desigualdad de Markov, tenemos que $P(Z(n,x) \geq 1) \geq E(Z(n,x))$; pero por comparación de series tenemos,

$$\sum_{n=0}^{\infty} P(Z(n,x) \geq 1) < \infty$$

por Borel-Cantelli se sigue,

$$P(lim\ supZ(n,x) \geq 1) = 0 \tag{5.10}$$

que se reduce a

$$P(Lim_{n\longrightarrow\infty}Z(n,x)=0)=1 \tag{5.11}$$

Para la demostración de (iii), consideremos una n^* con la propiedad de que,

$$E(Z(n,x))=b^{n^*}P(w_1+...+w_n\geq 0)>1$$

Esto se puede porque $P(w_1+...+w_n\geq 0)=e^{\ln(b)+l(0)}$. Construyamos un proceso de ramificación $K(k)$ de la siguiente manera. Los descendientes en la primera generación son los miembros del proceso de ramificación aleatorio que comenzando en cero, $x_0=0$, son positivos en la generación n^*, así las cosas, $K(1)=Z(n^*,0)$, de este modo la distribución del proceso de ramificación para $K(k)$ tiene distribución $Z(n^*,0)$. Así, hemos colocado la esperanza de este proceso de ramificación de tal manera que es mayor estricto que 1. Por el teorema 12 sabemos que la probabilidad de que $K(k)\longrightarrow\infty$ conforme k tiende a infinito es mayor estricto que cero. Ahora, construyamos K de tal modo que $K(k)\leq Z(kn^*,0)$, esto se logra ignorando de $x(kn^*,i)$ los valores negativos y adicionalmete los que son menores que el valor de alguno de sus ancestros, que es lo mismo a pedir que sea mayor que el de su progenitor, podemos sustituir $x_0=0$ por $x_0=y$ y la palabra positivos por mayores que y. Esto garantiza que $K(k)\leq Z(kn^*,0)$.

Como $Z(n,x)\geq Z(n,0)$ para toda $x\geq 0$, podemos concluir que la probabilidad de que $Z(n,x)\longrightarrow\infty$ a lo largo de la sucesión kn^* es positiva

Obsérvese cómo el resultado anterior sólo depende de b, no de la distribución del número de hijos. Sin embargo, para el cálculo de $l(0)$ se toma en cuenta qué tan alejado estoy de $E(w_n)$ y la distribución de w_n. El siguiente resultado resulta inmediato,

Corolario 2 *El teorema de Biggins y el lema 2 siguen siendo válidos si sustituimos $Z(n,x)$ por el número de individuos de la generación n que son estrictamente positivos.*

Para el Lema 2 obsérvese que la propiedad que estamos sustituyendo en este corolario no se utilizó en la demostración. Y para el teorema de Biggins hay que recordar que el teorema principal es válido para conjuntos de la forma (a,∞), en particular para $x=0$.

Una variación interesante la tenemos cuando nuestra población tiene una cierta habilidad (por ejemplo la fuerza, la velocidad con que corre, etc) y los únicos individuos que se reproducen son los que el valor de esta habilidad no es negativa. El número de hijos de un individuo que se puede reproducir no depende del valor de su habilidad. Queda claro que la extinción de los descendientes de un individuo no sólo depende de la esperanza del número de hijos, también que a lo largo de las generaciones siempre existe alguien que tenga la propiedad positiva. Tenemos un proceso de ramificación con barrera $z(n,x)$ que se obtiene de $Z(n,x)$ al eliminar los miembros que tienen a alguno de sus ancestros con su valor por debajo de cero. Formalmente

$$z(n,x) = \sum_{i=1}^{Y_n} 1_{\{x(j,L(j,n,i))>0 \,\forall\, j \in 0,\ldots n\}}$$

Teorema 14 *Consideremos el caso $b > 1$ y $E(w_n) < 0$*

(i) Para cualquier x,

$$Lim_{n \longrightarrow \infty} \frac{1}{n} \ln\left(E(z(n,x))\right) = \ln(b) - l(0).$$

(ii) Si $\ln(b) - l(0) < 0$, entonces $P(Lim_{n \longrightarrow \infty} z(n,x) = 0) = 1$

(iii) Si $\ln(b) - l(0) > 0$, entonces $P(Lim_{n \longrightarrow \infty} \sup z(n,x) = \infty) = 1$

El teorema anterior nos dice que la extinción de una población depende de dos cosas, algo que depende unicamente del número de hijos de un individuo que se puede reproducir, en particular de que tenga valor positivo y de $l(0)$, su valor depende de las variables aleatorias w_n que están directamente relacionadas con el valor de los individuos en la generación n.

Omitimos su demostración por ser similar a la del teorema 13; sin embargo en está requerimos del teorema 10.

Bibliografia

[1] Bartle, Robert G., Introducción al análisis matemático, Limusa, 1976

[2] Bazaraa, Mokhtar S. y C. M. Shetty, Nonlinear Programming, John Wiley & SONS, 1979.

[3] Biggins, Boris D. Lubachevsky, Adam Shwartz y Alan Weiss, A Branching radom walk with a Barrier, The Annals of Applied Probabliity Vol 1 No 4, $pp573 - 581$, 1991.

[4] Billingsley, Patrick, Probability and Measure, John Wiley & sons, 1995.

[5] Burkill, J. C. y H. Burkill, A Second course in Mathematical Analysis, Cambridge Mathematical Library, 1970.

[6] Den Hollander, Frank, Large Deviations, American Mathematical Society,1996.

[7] Dembo, Amir y Zeitouni, Large Deviations Techniques and Applications, Springer-Verlag, 1991.

[8] Eichelsbacher, Peter, Wahrscheinlichkeitstheorie, Vorlesungsskripte ($http : //www.ruhr - uni - bochum.de/imperia/md/content/stochastik/skripte/wtheorie_w s04.pdf$), 2004.

[9] Elstrodt, Jürgen , Mass- und Integrtionstheorie, Springer-Verlag, 2007, 5 edición.

[10] Feller, William , An Introduction to Probability and its Applications, John Wiley & SONS, 1957.

[11] Grabinsky Steider, Guillermo, notas del Curso de Análisis Matemático III, 2008.

[12] Meintrup, David y Stefan Schäffler, Stochastik Theorie und Anwendungen, Springer-Verlag, 2005.

[13]Ross, Sheldon M, Introduction Probability Models, Academic Press, 2007.

[14] Ross, Sheldon M. y Erol A. Pekóz, A Second Course in Probability, www.ProbabilityBookstore.com, 2007.

[15] Shwartz, Adam y Alan Weiss, Large deviations for performance analisys, Chapman & Hall, 1995.