56
2EJ

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

LA EXPANSION EN SERIES DE EDGEWORTH: UNA APLICACION A LA TEORIA DE RIESGO

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

GABRIELA MARTINEZ FENTANEZ

MEXICO, D. F. 1995

Universidad Nacional Autónoma de México

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

Comunicamos a usted que hemos revisado el trabajo de Tesis:
La Expansión en Series de Edgeworth: Una aplicación a la Teoría de Riesgo.

realizado por Gabriela Martínez Fentanez

con número de cuenta 8752267-9, pasante de la carrera de Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	M en C. Dino Mon Vásquez
Propietario	Mat. Jorge Francisco de la Vega Góngora
Propietario	Mat. Hugo Villaseñor Hernández
Suplente	Act. Noé Moacyr Vallejo González
Suplente	Act. Javier Ibarra Piña

Consejo Departamental de Matemáticas
M en C Alejandro Bravo Mojica

A Dios, a mi mami y a Paco

Con especial cariño a Tón y a los abuelos.

Gracias:

A mi mami, por toda la dedicación y el cariño con el que me ha guiado.

A Paco, por colmarme de felicidad y ánimo para buscar siempre nuevos objetivos que cumplir.

A Tón, por su grandiosa compañía.

A La Chatita y El Abuelo, por todo el amor que han puesto en mí.

A Mayo y Armando, por acogerme como unos buenos padres.

A la familia Asturiano Vera, por su cariño y confianza.

A Dino, a Jorge y a Claudia por su valiosa ayuda en este trabajo.

Al clan de actuarios de la Facultad de Ciencias, por todos los buenos momentos que hemos compartido.

A los amigos que afortunadamente he encontrado en la Dirección Técnica Vida.

A la Facultad de Ciencias de la U.N.A.M. y a todos aquellos profesores que contribuyeron a engrandecer mi conocimiento.

A todos aquellos que directa o indirectamente han contribuido a la realización de esta meta.

La realidad tiene una dimensión mágica que permite ponerle color para que el paso por esta vida no resulte tan aburrido.
I. Allende.

Contenido.

Introducción.

La necesidad de predecir la ocurrencia de eventos cuyas magnitudes presentan gran variabilidad, dio pie a un nuevo desarrollo dentro de la probabilidad, conocido como: "Teoria de Riesgo".

Dicha teoría ha tenido una gran aplicación en el campo de los seguros debido a las desviaciones de los valores esperados con respecto a los valores reales cuando se hacen predicciones acerca de la siniestralidad, del monto de reclamación agregada, del capital mínimo de garantía, etc. En la mayoría de los casos estas desviaciones son causadas por una inadecuada estimación, ya que las variables aleatorias que intervienen en cualquier proceso estocástico que se intenta describir son reemplazadas por simples promedios.

En este trabajo se presenta un panorama de la aplicación de la teoría de riesgo en el campo de los seguros, donde se pretende resolver uno de los problemas más comunes al que la técnica actuarial tiene que enfrentarse: "*La estimación de los montos reclamados por ocurrencia de siniestros*". Para cumplir con éste objetivo fue necesario investigar y desarrollar desde conceptos básicos, hasta aquellos más complejos, de los cuales algunos fueron analizados dentro de las materias de probabilidad y estadística en nuestro paso por la facultad de ciencias y otros son completamente nuevos, por tal motivo esta tesis puede ser útil a todos aquellos que deseen aprender o recordar conceptos importantes de la probabilidad y la estadística

En el capítulo 1, "Conceptos de límite", se presentan herramientas tales como lo son las definiciones y teoremas de convergencia, la función característica, el lema de Borel-Cantelli y la desigualdad de Tchebyshev, necesarias para comprender las técnicas de aproximación asintótica del capítulo 3.

En el capítulo 2, "Conceptos básicos y aplicaciones", se proporciona una amplia explicación acerca de sumas de variables aleatorias y sumas aleatorias de variables aleatorias, utilizando los momentos y cumulantes como medidas descriptivas de las mismas. Se establece la relación entre los momentos y los cumulantes, analizando las ventajas de trabajar con estos últimos, se introduce el concepto de cumulante estandarizado y su orden de convergencia, indispensable para el desarrollo de las expansiones en serie de Edgeworth del capítulo 3. Para finalizar este capítulo se proponen técnicas a través del proceso poisson compuesto para calcular la reclamación agregada en una cartera de seguros, presentando como inconvenientes, los sesgos en los resultados obtenidos, razón por la que se recomienda utilizar otros métodos para calcular dicha reclamación agregada, tales como las técnicas asintóticas presentadas en el capítulo 3.

Para concluir, en el capítulo 3 se analizan y comparan las técnicas de aproximación asintótica, como la ley de los grandes números, el teorema del límite central y la expansión en series de Edgeworth, demostrando que esta última presenta menos desviaciones que las otras dos, ya que por ser una expresión formada por potencias de distribuciones normales estandarizadas, corrige los errores de una simple normal estándar. Adicionalmente para los fines de este trabajo, se utilizó la inversión de Cornish-Fisher, misma que se aplica a las expansiones en series de Edgeworth, haciendo posible el cálculo de los cuantiles de una probabilidad. Para concretar la idea principal de este trabajo se desarrolló un ejemplo práctico tomado de los problemas frecuentes de una compañía de seguros, con el fin de probar la estrecha relación entre los seguros y la teoría de la probabilidad y estadística.

Capítulo 1

Conceptos de Límite.

1.1 Convergencia.

Uno de los conceptos más usados en estadística es el de *límite*. Ejemplos de ello se pueden encontrar en la propiedad de consistencia de los estimadores; en la distribución asintótica de los estimadores máximo verosímiles e incluso en la elección del tamaño de una muestra que satisfaga ciertos compromisos de precisión y confianza. En la estadística, son de interés dos tipos de aproximaciones, entre variables aleatorias y entre funciones de distribución. En esta sección se revisarán las definiciones de estos tipos de convergencia y sus relaciones, que servirán como base para los teoremas de convergencia que se tratarán en este trabajo.

1.1.1 Convergencia Casi Segura, en Probabilidad, en R-media y en Distribución.

Supóngase que $\{Y_n : n \in \mathbf{N}\}$ es una sucesión de variables aleatorias independientes (v.a.i.) con distribución *Bernoulli* con parámetro θ, y sea $S_n = Y_1 + Y_2 + \cdots + Y_n$, entonces

$$P(S_n = k) = \binom{n}{k} \theta^k (1-\theta)^{n-k} \text{ para toda } k \in J_n^0, \tag{1.1.1}$$

donde J_n^0 es un conjunto de índices para $k = 1 \dots n$.

Debe notarse que cuando n es grande el cálculo se vuelve complicado, por tal razón es factible preguntarse si para estos casos (1.1.1) puede aproximarse a $\{P_k : k \in \mathbf{N}\}$ que sea más fácil de calcular, es decir, se desea conocer

$$\lim_{n\to\infty} P(S_n = k) = P_k \text{ para toda } k \in \mathbf{N}.$$

Por otro lado se considera ahora a $X_n = \frac{1}{n}S_n$ para toda $n \in \mathbf{N}$ donde,

$$E(X_n) = E\left(\frac{S_n}{n}\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} Y_i\right) = \theta,$$

lo que significa que X_n es un estimador insesgado de θ, pero no se sabe qué tanto se parece a θ. Como la varianza de X_n, es

$$Var(X_n) = \frac{1}{n^2}Var(S_n) = \frac{\theta(1-\theta)}{n}, \tag{1.1.2}$$

se sigue que $Var(X_n) \to 0$ cuando $n \to \infty$, en este caso se tiene que

$$\lim_{n\to\infty} E(X_n - \theta)^2 = 0. \tag{1.1.3}$$

Intuitivamente (1.1.3) establece que X_n se parece cada vez más a θ a medida que n crece; pero no se puede decir con seguridad que en el límite lleguen a ser iguales, es decir, que $\lim_{n\to\infty} X_n = \theta$. (Ver figura 1.1.1)
Por otro lado, si $\varepsilon \in \mathbf{R}^+$, puede calcularse

$$P(|X_n - \theta| \geq \varepsilon), \tag{1.1.4}$$

si esta probabilidad se hace pequeña a medida que n crece, podría decirse que X_n se acerca a θ si n crece, lo cual provocaría en algún sentido que

$$\lim_{n\to\infty} X_n = \theta. \tag{1.1.5}$$

De esta forma (1.1.3) y (1.1.4) parecen dos formas de decir lo mismo, pero no se conoce cual es la mejor o si son equivalentes a (1.1.5).

GRAFICA COMPARATIVA UTILIZANDO VALORES BERNOULLI (0.5) SIMULADOS

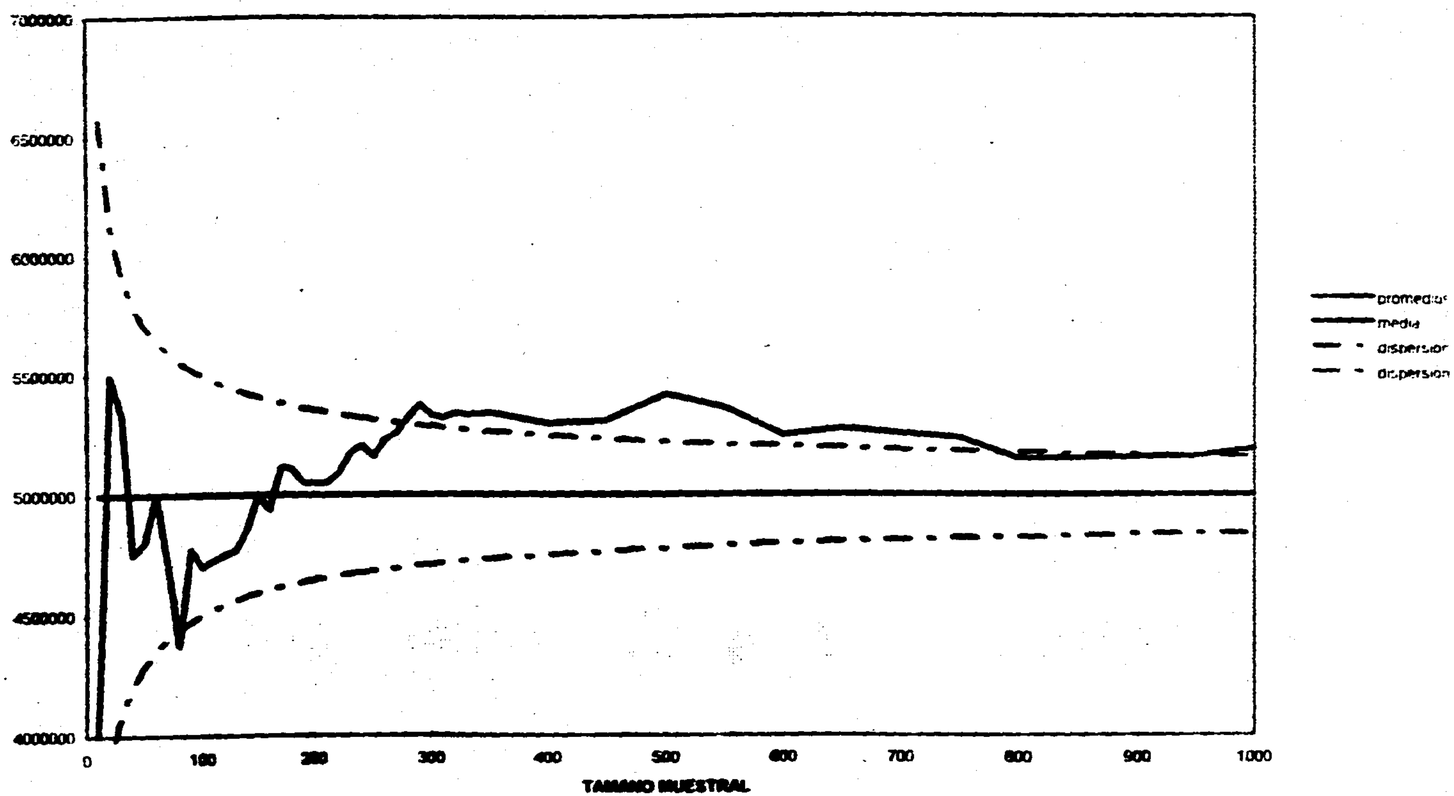


figura 1.1.1

A continuación se enunciarán una serie de resultados y definiciones que son ampliamente conocidos en la literatura de probabilidad.

Recordando que $\{Y_n; n \in \mathbf{N}\}$ es una sucesión de variables aleatorias definidas sobre $(\Omega, \mathcal{F}, \mathcal{P})$ y Y otra variable aleatoria definida sobre el mismo espacio, se tiene la siguiente definición.

Definición 1.1.1 (Convergencia Puntual.) *El* $\lim_{n\to\infty} Y_n = Y$ *en* Ω *si para toda* $\varepsilon \in \mathbf{R}^+$ *y* $w \in \Omega$ *existe* $n_{\varepsilon,w} \in \mathbf{N}$ *tal que*

$$|Y_n(w) - Y(w)| < \varepsilon \text{ para toda } n \geq n_{\varepsilon,w}$$

denotándolo por $Y_n \to Y$ *en* Ω.

Si $n_{\varepsilon,w}$ no depende de w, se dice que Y_n converge uniformemente a Y en Ω denotándolo por $Y_n \xrightarrow{u} Y$ en Ω. Además si $Y_n \xrightarrow{u} Y$ en Ω entonces $Y_n \to Y$ en Ω.

Si en (1.1.4) se tiene que para toda $\varepsilon \in \mathbf{R}^+$

$$P[|X_n - \theta| \geq \varepsilon] \to 0 \text{ si } n \to \infty, \tag{1.1.6}$$

esto podría implicar que $X_n \to \theta$ en Ω. Si $A_n = \{w \in \Omega : |X_n(w) - \theta| \geq \varepsilon\}$, la expresión anterior puede escribirse como $\lim_{n\to\infty} P(A_n) = 0$, así que lo que establece (1.1.6) es que la probabilidad de que $X_n(w)$ difiera de θ en más de ε se acerca a cero si n crece, pero puede ocurrir que para n grande X_n y θ difieran en más de ε. Por lo tanto la convergencia en (1.1.6) es más débil que la convergencia puntual y que la uniforme en Ω lo que motiva a pensar en algún otro tipo de convergencia que de alguna forma refleje más fuerza en el modo de acercarse al límite. Si se toma el siguiente conjunto

$$B_n = \{w \in \Omega : \lim_{n\to\infty} X_n(w) = X(w)\}$$

y se obtiene que $P(B_n) = 1$, esto necesariamente reflejaría una condición de convergencia mucho más fuerte a la expresada en (1.1.6), ya que no se habla de que en el límite el conjunto de las w's que hacen que $X_n(w)$ se aleje de $X(w)$ tiene probabilida cero, sino que se habla de la medida del conjunto de las w's para las cuales existe la convergencia de la variable $X_n(w)$ a $X(w)$.

Además de lo anterior, también puede ocurrir, como en cálculo, que la sucesión, (que en este caso es de v.a.), X_n no converja a θ para algún $w \in \Omega$; por ejemplo, supóngase que $w \in \Omega$ es tal que $X_n(w) = 1$ para toda $n \in \mathbf{N}$ y que $\theta \in (0, 1)$.

Definición 1.1.2 (Convergencia Casi Segura) *Sea $\{Y_n : n \in \mathbf{N}\}$ y Y variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$, entonces, se dice que Y_n converge casi seguramente, (o con probabilidad 1) a Y si existe $A \in \mathcal{F}$ nulo (i.e. $P(A) = 0$) tal que $Y_n(w) \to Y(w)$ para toda $w \in A^c$ y se denota por $Y_n \xrightarrow{c.s.} Y$. Esto es equivalente a decir que*

$$P\{w \in \Omega; \lim_{n\to\infty} Y_n(w) = Y(w)\} = 1.$$

Si $A = \emptyset$, se tiene convergencia puntual.

Definición 1.1.3 (Convergencia en Probabilidad.) *Sea $\{Y_n\}$ y Y variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$; entonces,se dice que Y_n converge en probabilidad a Y si*

$$\lim_{n\to\infty} P\{w \in \Omega : |Y_n(w) - Y(w)| \geq \epsilon\} = 0 \text{ para toda } \epsilon \in \mathbf{R}^+,$$

y se denota por $Y_n \xrightarrow{P} Y$.

Definición 1.1.4 (Convergencia en R-media) *Sea $\{Y_n\}$ y Y variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$; entonces, se dice que Y_n converge en r-media a Y si*

$$\lim_{n\to\infty} E(|Y_n - Y|^r) = 0, \; r \geq 1.$$

y se denota por $Y_n \xrightarrow{r} Y$.

En esta definición se está suponiendo implícitamente que $E(|Y_n|^r) < \infty$ para toda $n \in \mathbf{N}$ y $E(|Y|) < \infty$.

Definición 1.1.5 (Convergencia en Distribución.) *Sea $\{Y_n\}$ y Y variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$; entonces, si $\{F_n\}$ y F son las funciones de distribución de $\{Y_n\}$ y Y respectivamente, se dice que Y_n converge en distribución o débilmente a Y si*

$$\lim_{n\to\infty} F_n(y) = F(y)$$

para toda y punto de continuidad de F, denotándose como $Y_n \xrightarrow{d} Y$.

Debido a que puede definirse la función de distribución sin necesidad de hablar de variables aleatorias, la definición anterior puede darse en términos de $\{F_n\}$ y de F, en cuyo caso se dirá que la sucesión $\{F_n\}$ converge débilmente a F, y se denota como $F_n \xrightarrow{d} F$.

Las relaciones entre estos tipos de convergencia son las que se enuncian en el siguiente Teorema.

Teorema 1.1.1 *La convergencia puntual implica la convergencia casi segura y la convergencia en r-media; de la convergencia casi segura se sigue la convergencia en probabilidad; la convergencia en r-media implica convergencia en probabilidad y la convergencia en probabilidad implica la convergencia en distribución.*

Otra caracterización de convergencia casi segura es:

Lema 1.1.1

$X_n \xrightarrow{c.s} X$ *si y sólo si* $\lim_{n\to\infty} P[w \in \Omega : |X_m(w) - X(w)| < \varepsilon$ *para toda* $m \geq n] = 1$

para toda $\varepsilon \in \mathbb{R}^+$

Demostración.

Sea $C = \{w \in \Omega : \lim_{n\to\infty} X_n(w) = X(w)\}$ y para cada $\varepsilon \in \mathbf{R}^+$ y $n, k \in \mathbf{N}$, defínase a

$$A_{n,\varepsilon} = \{w \in \Omega : |X_m(w) - X(w)| < \varepsilon \text{ para toda } m \geq n\}$$

y

$$B_{n,k} = \{w \in \Omega : |X_m(w) - X(w)| < \frac{1}{k} \text{ para toda } m \geq n\},$$

entonces

$$\lim_{\varepsilon\to 0} \bigcup_{n\in\mathbf{N}} A_{n,\varepsilon} = \bigcap_{k\in\mathbf{N}} \bigcup_{n\in\mathbf{N}} B_{n,k} = C, \tag{1.1.7}$$

como $B_{n,k} \subseteq B_{(n+1),k}$ para toda $n \in \mathbf{N}$ y P es continua, se sigue que

$$P(C) = P\left(\bigcap_{k\in\mathbf{N}} \bigcup_{n\in\mathbf{N}} B_{n,k}\right) = \lim_{k\to\infty} \lim_{n\to\infty} P(B_{n,k}) \tag{1.1.8}$$

y

$$P(C) \leq P\left(\bigcup_{n\in\mathbf{N}} B_{n,k}\right) = P\left(\bigcup_{n\in\mathbf{N}} A_{n,\varepsilon}\right) = \lim_{n\to\infty} P(A_{n,\varepsilon}). \tag{1.1.9}$$

Si $X_n \xrightarrow{c.s} X$ entonces $P(C) = 1$ y (1.1.8) implica que,

$$\lim_{n\to\infty} P(A_{n,\varepsilon}) = 1$$

para toda $\varepsilon \in \mathbf{R}^+$. Análogamente, si $\lim_{n\to\infty} P(A_{n,\varepsilon}) = 1$ para toda $\varepsilon \in \mathbf{R}^+$ (1.1.7) y (1.1.9) implican que $P(C) = 1$ estableciendo así la equivalencia.

□

Para aclarar las definiciones de convergencia citadas anteriormente y el *Teorema 1.1.1* se procederá a dar algunos ejemplos.

Ejemplo 1.1.1

Sea $\Omega = \{w_1, w_2, w_3, w_4\}$, $\mathcal{F} = 2^\Omega$, $P(\{w_j\}) = \frac{1}{4}$ para toda $j \in J_4$.

Defínase a $\{Y_n\}$ y a Y de la siguiente forma.

$$\begin{aligned} Y_n(w_1) = Y_n(w_2) = Y(w_3) = Y(w_4) &= 1 \\ Y_n(w_3) = Y_n(w_4) = Y(w_1) = Y(w_2) &= 0 \end{aligned}$$

para toda $n \in \mathbf{N}$

Si $\{F_n; n \in \mathbf{N}\}$ y F son las funciones de distribución correspondientes, se puede ver que

$$F(y) = F_n(y) = \begin{cases} 0 & \text{si } Y < 0 \\ \frac{1}{2} & \text{si } 0 \leq Y \leq 1 \\ 1 & \text{si } Y > 1 \end{cases}$$

Por lo que $F(x) = F_n(x)$. Sin embargo $|Y_n(w) - Y(w)| = 1$ para toda $w \in \Omega$, entonces $P\{w \in \Omega : |Y_n(w) - Y(w)| \geq \varepsilon\} = 1$, por lo tanto Y_n no converge en probabilidad ni casi seguramente. Esto quiere decir que convergencia en distribución no implica convergencia en probabilidad ni casi segura.

Además

$$E(|Y_n - Y|^r) = 1 \text{ para toda } r \in \mathbf{N},$$

por lo que puede decirse que Y_n no converge en r-media a Y.

Ejemplo 1.1.2

Sea $X \sim Blli(1/2)$, entonces $Y = 1 - X \sim Blli(1/2)$. Si $X_n = X$ para toda $n \in \mathbf{N}$ se cumple que $X_n \xrightarrow{d} Y$, pero $|X_n - Y| = 1$ para toda $n \in \mathbf{N}$ esto implica que $P[|X_n - Y| \geq \varepsilon] = 1$ para toda $\varepsilon \in (0, 1)$ y entonces X_n no converge en probabilidad a Y. También aquí se observa que convergencia en distribución no implica convergencia en probabilidad.

Ejemplo 1.1.3

Sea $\{X_n; n \in \mathbf{N}\}$ una sucesión de de v.a.i. con distribución *Bernoulli* con $P(X_n = 1) = \frac{1}{n}$ para toda $n \in \mathbf{N}$.

Como $P[X_n \neq 0] = \frac{1}{n}$, entonces

$$\lim_{n\to\infty} P[X_n \neq 0] = 0$$

y

$$\lim_{n\to\infty} P[|X_n| \geq \varepsilon] = 0.$$

Por lo tanto

$$X_n \xrightarrow{P} 0.$$

Por otro lado sea $\varepsilon \in (0,1)$ y defínase

$$B_{n,\varepsilon} = \{w \in \Omega : |X_m(w)| < \varepsilon \ \forall\, m \geq n\} \text{ para toda } n \in \mathbf{N}$$

entonces

$$P(B_{n,\varepsilon}) = \lim_{M\to\infty} P[|X_m| < \varepsilon,\ n \leq m \leq M]$$

pues P es continua. Lo que implica que

$$P(B_{n,\varepsilon}) = \lim_{M\to\infty} P[X_n = 0, X_{n+1} = 0, \ldots, X_M = 0],$$

y por independencia

$$\begin{aligned} P(B_{n,\varepsilon}) &= \lim_{M\to\infty} \left(1 - \frac{1}{n^r}\right)\left(1 - \frac{1}{n^r+1}\right)\cdots\left(1 - \frac{1}{M}\right) \\ &= \lim_{M\to\infty} \left(\frac{n-1}{n}\right)\left(\frac{n}{n+1}\right)\left(\frac{n+1}{n+2}\right)\cdots\left(\frac{M-1}{M}\right) \\ &= \lim_{M\to\infty} \frac{n-1}{M} = 0, \end{aligned}$$

lo que implica que X_n no converge casi seguramente a cero, pero $X_n \xrightarrow{r} 0$, pues

$$E(|X_n|) = 0 \cdot P(X_n = 0) + 1 \cdot P(X_n = 1) = \frac{1}{n}$$

lo que implica que $\lim_{n\to\infty} E[|X_n|^r] = 0$, para toda n, $r \in \mathbf{N}$.

De aquí que convergencia en probabilidad no implica convergencia casi segura. También se puede ver que convergencia en r-media no implica convergencia casi segura.

Ejemplo 1.1.4

Sea $(\Omega, \mathcal{F}, P) = ((0,1], \mathbf{B}_{(0,1]}, \lambda_{(0,1]})$ y sea $\{X_n : n \in \mathbf{N}\}$ la sucesión definida por

$$\begin{array}{lcl} X_1 = I_{(0,\frac{1}{2}]} & , & X_2 = I_{(\frac{1}{2},1]}, \\ X_3 = I_{(0,\frac{1}{3}]} & , & X_4 = I_{(\frac{1}{3},\frac{2}{3}]}, \\ X_5 = I_{(\frac{2}{3},1]} & , & X_6 = I_{(0,\frac{1}{4}]}, \ldots \end{array}$$

Con $\mathbf{B}_{(0,1]}$ = los borelianos en $(0,1]$ y $\lambda_{(0,1]}$ = la medida de Lebegue sobre los intervalos $(0,1]$.
Puede observarse que $E[|X_n|^r] \to 0$ por lo tanto $X_n \xrightarrow{P} 0$, pero X_n no converge casi seguramente a 0.

Ejemplo 1.1.5

Sea $(\Omega, \mathcal{F}, P)$ como el ejemplo anterior, y defínase a $X_n = nI_{(0,\frac{1}{n})}$ para toda $n \in \mathbf{N}$ entonces

$$A_n = \left\{w \in \Omega : |X_n(w)| < \epsilon\right\} = \left\{w \in \Omega : X_n(w) \in \left(\frac{1}{n}, 1\right)\right\},$$

y como $A_n \subset A_{n+1}$ para toda $n \in \mathbf{N}$ se sigue que

$$\begin{aligned} \lim_{n\to\infty} P(A_n) &= P\left(\lim_{n\to\infty} A_n\right) \\ &= P(\Omega) = 1 \text{ para toda } \epsilon \in \mathbf{R}^+, \end{aligned}$$

por lo que $X_n \xrightarrow{c.s.} 0$. Pero $E(X_n) = n\left(\frac{1}{n}\right) = 1$ para toda $n \in \mathbf{N}$ y entonces X_n no converge en media a cero. Por lo tanto convergencia casi segura no implica convergencia en r-media.

Ejemplo 1.1.6

Sea $X_n = \frac{1}{n}$ para toda $n \in \mathbf{N}$ y $X = 0$, es claro que $X_n \xrightarrow{c.s.} 0$.

Defínase a

$$F_n(x) = \begin{cases} 1 & \text{si } x \geq \frac{1}{n} \\ 0 & \text{si } x < \frac{1}{n} \end{cases}$$

y

$$F(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

entonces $F_n(x) \to F(x)$ para toda $x \neq 0$ punto de continuidad de la distribución límite pues $F_n(0) = 0 \neq F(0) = 1$, como 0 no está en C_F, entoces $F_n \xrightarrow{d} F$.

En general la convergencia en distribución no proporciona la información acerca del comportamiento límite de las variables aleatorias como se puede ver en el siguiente ejemplo:

Ejemplo 1.1.7

Sea $X_n(w) = (-1)^n X(w)$ para $n \geq 1$, donde $X(\cdot)$ es una v.a. $N(0,1)$. Claramente $X_n \sim N(0,1)$ para toda $n \geq 1$, por lo tanto converge en distribución.

Por otro lado $X_n(w)$ diverge excepto en $X(w) = 0$ y esto ocurre con probabilidad cero, de aquí que convergen las distribuciones pero no las variables aleatorias.

Es bastante difícil demostrar con las funciones de distribución la convergencia. Un método que en muchos casos simplifica los procedimientos se basa en el *Teorema de continuidad* para funciones características que se enunciará en la siguiente sección, pero antes será necesario hacer un estudio breve acerca de las funciones características ya que este teorema se basa en ellas.

1.2 Función Característica.

Para entender esta función y sus propiedades es necesario recordar conceptos de variable compleja.

1. Si $z \in C$ y $z = x + iy$, entonces $|z| = |x^2 + y^2| = \sqrt{x^2 + y^2}$

2. La distancia entre dos números complejos z_1 y z_2 es $|z_2 - z_1|$.

3. Si una función real de variable real tiene una expansión en series de potencias con radio de convergencia positivo, se puede utilizar la misma serie para definir una función compleja, por ejemplo,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!},$$

para cualquier número complejo z. También la relación $e^{z_1+z_2} = e^{z_1}e^{z_2}$ permanece válida cuando z_1 y $z_2 \in C$

4. Sea $Z = it$, donde $t \in \mathbb{R}$, entonces

$$\begin{aligned} e^{it} &= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \\ &= 1 + it - \frac{t^2}{2} - \frac{it^3}{3!} + \frac{t^4}{4!} + \cdots \\ &= \left(1 - \frac{t^2}{2} + \frac{t^4}{4} - \cdots\right) + i\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots\right) \end{aligned}$$

Estas series de potencias corresponden a las expansiones de $\cos(t)$ y $\sin(t)$, por lo que:

$$e^{it} = \cos t + i \sin t,$$

dado que $\cos(-t) = \cos t$ y $\sin(-t) = -\sin t$, entonces

$$e^{-it} = \cos t - i \sin t,$$

por lo que

$$\cos t = \frac{e^{it} + e^{-it}}{2} \quad \sin t = \frac{e^{it} - e^{-it}}{2i}$$

y además

$$\left|e^{it}\right| = \left(\cos^2 t + \sin^2 t\right)^{1/2} = 1.$$

5. Si $f(t)$ y $g(t)$ son funciones reales de variable real, entonces

$$h(t) = f(t) + ig(t),$$

es una función compleja. Además se puede diferenciar $h(t)$, es decir

$$h'(t) = f'(t) + ig'(t)$$

si f' y $g'(t)$ existen, es decir la i se toma como constante. Similarmete se define

$$\int_a^b h(t)dt = \int_a^b f(t)dt + i\int_a^b g(t)dt,$$

siempre que existan las integrales.

6. La fórmula

$$\frac{d}{dt}e^{ct} = ce^{ct},$$

es válida para cualquier constante compleja.

7. *El Teorema Fundamental del Cálculo* se sigue cumpliendo si c es una constante compleja, es decir

$$\int_a^b e^{ct}dt = \frac{e^{cb} - e^{ca}}{c}$$

8. Una v.a. compleja Z puede ser escrita en la forma $Z = X + iY$, donde X y Y son v.a. reales y su esperanza está dada por

$$E(Z) = E(X) + iE(Y),$$

siempre que $E(X)$ y $E(Y)$ existan. Al igual que en el caso de variables reales, $E(Z)$ existe si $E(|Z|)$ existe y en este caso $|E(Z)| \leq E(|Z|)$.

9. La fórmula

$$E(a_1Z_1 + a_2Z_2) = a_1E(Z_1) + a_2E(Z_2)$$

es válida siempre que a_1 y a_2 sean constantes complejas y Z_1 y Z_2 sean v.a. complejas con esperanza finita.

10. Suponga que X es una v.a. y $t \in R$, entonces por (4), $|e^{itX}| = 1$, por lo tanto $E(e^{itX})$ siempre existe, por lo que la siguiente definición tiene sentido.

Definición 1.2.1 *Sea X una v.a. real, defínase a la función característica $\varphi_X(t)$ de X por*

$$\varphi_X(t) = E(e^{itX}) \ , \ t \in \mathbb{R}.$$

Además de los puntos anteriores acerca de la función característica, es importante mencionar que la razón por la que la función generadora de momentos no siempre existe y la característica siempre existe es que e^{tX} no siempre está acotada y e^{itX} siempre está acotada.

Una vez hecho este recordatorio de la función característica se procederá a dar algunos ejemplos

Ejemplo 1.2.1

1. Sea X una v.a. tal que $P(X = a) = 1$, entonces $\varphi_X(t) = E(e^{itX}) = \sum_x e^{itX} P(X = x) = e^{ita}$ para $t \in \mathbb{R}$. En particular si X toma el valor de cero su función característica es igual a uno.

2. Sea $U \sim U(a,b)$, entonces para $t \neq 0$ se tiene

$$\begin{aligned} \varphi_U(t) &= E\left(e^{itU}\right) = \int_a^b \frac{e^{itu}}{b-a} du \\ &= \left[\frac{e^{itu}}{it(b-a)}\right]_a^b \\ &= \frac{e^{itb} - e^{ita}}{it(b-a)} \end{aligned}$$

en particular si $U \sim U(-1,1)$ se tiene que

$$\varphi_U(t) = \frac{1}{2}\left(\frac{e^{it} - e^{-it}}{it}\right) = \frac{\sin t}{t}$$

3. Si $X \sim Poisson(\lambda)$ entonces,

$$\begin{aligned}
\varphi_X(t) &= E(e^{itX}) = \sum_{x=0}^{\infty} \frac{e^{itx}\lambda^x e^{-x}}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{it})^x}{x!} \\
&= \exp\left\{-\lambda(1-e^{it})\right\}
\end{aligned}$$

4. Sea $X \sim Exp(\lambda)$, entonces

$$\begin{aligned}
\varphi_X(t) &= \int_0^{\infty} \lambda e^{itx} e^{-\lambda x} dx \\
&= \lambda \int_0^{\infty} e^{-x(\lambda - it)} dx \\
&= \left[\frac{-\lambda}{\lambda - it} e^{-x(\lambda - it)}\right]_0^{\infty} \\
&= \left(\frac{\lambda}{\lambda - it}\right)
\end{aligned}$$

5. Si $X \sim Bin(n,p)$, entonces

$$\begin{aligned}
\varphi_X(t) &= E(e^{itX}) = \sum_{x=0}^{n} e^{itx} \binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (pe^{it})^x q^{n-x} \\
&= (pe^{it} + q)^n
\end{aligned}$$

Teorema 1.2.1 *Si X es una v.a. y a, b son constantes entonces*

$$\varphi_{a+bX}(t) = e^{ita}\varphi_X(tb).$$

Demostración.

$$\begin{aligned}
\varphi_{a+bX}(t) &= E\left(e^{it(a+bX)}\right) \\
&= E\left(e^{ita}e^{itbX}\right) \\
&= e^{ita}E\left(e^{itbX}\right) \\
&= e^{ita}\varphi_X(tb). \qquad (1.2.10)
\end{aligned}$$

□

Ejemplo 1.2.2

Si $X \sim U(a,b)$, por la proposición anterior, y debido a que

$$X = \underbrace{\frac{a+b}{2}}_{\text{desplazamiento}} + \underbrace{\frac{b-a}{2}}_{\text{magnitud}} U$$

donde $U \sim U(-1,1)$, ya que

$$\begin{aligned}
P\left[\frac{a+b}{2} + \frac{b-a}{2}U \leq x\right] &= P\left[\frac{b-a}{2}U \leq x - \frac{a+b}{2}\right] \\
&= P\left[U \leq \left(x - \frac{a+b}{2}\right)\frac{2}{b-a}\right] \\
&= \int_{-1}^{\left(x-\frac{a+b}{2}\right)\frac{2}{b-a}} \frac{1}{2}dt \\
&= \int_a^x \frac{1}{b-a}dt \\
&= F_{U'}(x) \text{ donde } U' \sim U(a,b)
\end{aligned}$$

y entonces la función característica es:

$$\varphi_X(t) = \exp\left\{it\left(\frac{a+b}{2}\right)\right\}\frac{\sin(b-a)(t/2)}{\frac{(b-a)t}{2}}.$$

Teorema 1.2.2 *Sean X y Y v.a.i., entonces*

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) \text{ para toda } t \in \mathbb{R}.$$

Demostración.

$$\begin{aligned} \varphi_{X+Y}(t) &= E\left(e^{it(X+Y)}\right) \\ &= E\left(e^{itX}e^{itY}\right) \\ &= \varphi_X(t)\varphi_Y(t). \end{aligned}$$

□

En general si $X_1, X_2, \ldots, X_n$ son v.a.i. y $S_n = \sum_{i=1}^{n} X_i$, entonces

$$\varphi_{S_n}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t).$$

Teorema 1.2.3 *Si X es una v.a. tal que tiene momento de orden k, entonces $\varphi_X(t)$ tiene derivada de orden k y está dada por*

$$\varphi_X^{(k)}(t) = \int_{-\infty}^{\infty} (ix)^k e^{itx} f(x)dx,$$

es decir se puede derivar bajo el signo de la integral y

$$\varphi_X^{(k)}(t) = E\left((iX)^k e^{itX}\right),$$

para $t = 0$

$$\varphi_X^{(k)}(0) = i^k \underbrace{\int_{-\infty}^{\infty} X^k f(x)dx}_{\text{momento } k \text{ de un v.a.}}$$

y finalmente

$$\varphi_X^{(k)}(0) = i^k E(X^k).$$

Recíprocamente si $\varphi_X(t)$ tiene derivada finita de orden par k en $t = 0$ entonces X tiene momento finito de orden k, y si k es impar X tiene momento finito de orden $k-1$.

Teorema 1.2.4 *Si X en una v.a. con momento absoluto de orden $k \geq 1$, entonces $\varphi_X(t)$ tiene las siguientes expansiones en una vecindad de cero.*

$$\varphi_X(t) = \sum_{j=0}^{k} \frac{i^j}{j!} m_j t^j + o(|t|^k)\ t \to 0$$

$$\varphi_X(t) = \sum_{j=0}^{k-1} \frac{i^j}{j!} m_j t^j + \frac{\theta_k}{k!} n_k |t|^k \quad |\theta_k| \leq 1.$$

donde

$$\begin{aligned} m_j &= E(X^j) = \int X^j f(x)dx\ j = 0, 1, \ldots, k \\ n_k &= E(|X|^k) = \int |X|^k f(x)dx. \end{aligned}$$

En particular si X es una v.a. tal que existen todos sus momentos, entonces

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{i^n E(X^n) t^n}{n!}$$

Ejemplo 1.2.3

Sea X una v.a. $N(0, \sigma^2)$, encontrar su función característica.

Solución.

Se sabe que $E(X^n) = 0$, si n es impar, ya que

$$\int_{-\infty}^{-\infty} x^{(2k-1)} f_X(x)dx = \int_{-\infty}^{0} x^{(2k-1)} f_X(x)dx + \int_{0}^{\infty} x^{(2k-1)} f_X(x)dx,$$

sea $y = x$, lo que implica que si $x = -\infty$ entonces $y = -\infty$, si $x = 0$ entonces $y = 0$ y $dy = dx$, por lo tanto

$$\begin{aligned} \int_{-\infty}^{\infty} x^{(2k-1)} f_X(x)dx &= -\int_{-\infty}^{0} (-y)^{2k-1} f_X(-y)dy + \int_{0}^{\infty} (x)^{(2k-1)} f_X(x)dx \\ &= 0, \end{aligned}$$

además, también se sabe que

$$E\left(X^{2k}\right) = \frac{\sigma^{2k}(2k)!}{2^k k!},$$

ya que si X se distribuye como $N(0,\sigma^2)$, entonces X^2 se distribuye como $\Gamma(\frac{1}{2}, \frac{1}{2\sigma^2})$, es decir el momento de orden k de la *Gamma* es el momento de orden $2k$ de la *Normal*. Se calculará ahora el momento m de una densidad $\Gamma(\alpha,\lambda)$.

$$\begin{aligned}
E(X^m) &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^m x^{\alpha-1}e^{-\lambda x}dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{(m+\alpha)-1}e^{-\lambda x}dx \\
&= \frac{\Gamma(m+\alpha)}{\Gamma(\alpha)\lambda^m} \\
&= \frac{1}{\lambda^m}\frac{\alpha(\alpha+1)\cdots(\alpha+m-1)\Gamma(\alpha)}{\Gamma(\alpha)} \\
&= \frac{(\alpha)(\alpha+1)\cdots(\alpha+m-1)}{\lambda^m} \\
&= \frac{\Gamma(m+\alpha)}{\lambda^m},
\end{aligned}$$

pero si $\alpha = \frac{1}{2}$ entonces $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ y

$$\begin{aligned}
\Gamma(\alpha+1) &= \int_0^\infty x^\alpha e^{-x}dx \\
&= -x^\alpha e^{-x}\mid_0^\infty + \int_0^\infty \alpha x^{\alpha-1}e^{-x}dx \\
&= \alpha\Gamma(\alpha) \\
\Gamma(\alpha+m) &= \int_0^\infty x^{\alpha+m-1}e^{-x}dx \\
&= -x^{\alpha+m-1}e^{-x}\mid_0^\infty + \int_0^\infty (\alpha+m-1)x^{\alpha+m-2}e^{-x}dx \\
&= (\alpha+m-1)\Gamma(\alpha+m-1) \\
&= (\alpha+m-1)(\alpha+m-2)\Gamma(\alpha+m-2) \\
&= (\alpha+m-1)(\alpha+m-2)(\alpha+m-3)\ldots(\alpha+1)(\alpha)\Gamma(\alpha),
\end{aligned}$$

dado que $2k$ de la *Normal* es el momento k de la Gamma $(\frac{1}{2}, \frac{1}{2\sigma^2})$, lo que implica que

$$E(X^{2k}) = E[(X^2)^k] = E[Y^k],$$

donde X se distribuye como *Normal* con parámetros $(0, \sigma^2)$ y Y se distribuye como *Gamma* con parámetros $(\frac{1}{2}, \frac{1}{2\sigma^2})$, entonces

$$\begin{aligned}
E(Y^k) &= \frac{\Gamma(k+\frac{1}{2})}{(\frac{1}{2\sigma^2})^k \Gamma(\frac{1}{2})} \\
&= \frac{\frac{1}{2}\frac{3}{2}\cdots\frac{2k-1}{2}}{(\frac{1}{\sigma^2})^k} \\
&= \frac{\frac{1}{2^k}\left[\sigma^{2k}(1\cdot 3\cdot 5\cdots 2(k-1))\right]}{\frac{1}{2^k}} \left[\frac{2\cdot 4\cdot 6\cdot 8\cdots 2k}{2\cdot 4\cdot 6\cdot 8\cdots 2k}\right] \\
&= \frac{\sigma^{2k}(2k)!}{(1\cdot 2\cdot 3\cdots k)2^k} \\
&= \frac{\sigma^{2k}(2k)!}{2^k k!} = E(X^{2k})
\end{aligned}$$

ahora

$$\begin{aligned}
\varphi_X(t) &= \sum_{k=0}^{\infty} \frac{i^{2k}t^{2k}}{(2k)!}\frac{\sigma^{2k}(2k)!}{2^k k!} \\
&= \sum_{k=0}^{\infty} \frac{(\frac{-\sigma^2 t^2}{2})^k}{k!} \\
&= \exp -\frac{\sigma^2 t^2}{2},
\end{aligned}$$

y como

$$\begin{aligned}
W \sim N(\mu, \sigma^2) \quad &\text{esto implica} \quad W = (W-\mu)+\mu \\
&\text{se tiene} \quad W-\mu \sim N(0,\sigma^2) \\
&\text{se tiene} \quad \varphi_W(t) = e^{it\mu}e^{-\sigma^2 t^2/2} \\
&\text{y así} \quad \varphi_W(t) = \exp\left\{it\mu - \frac{\sigma^2 t^2}{2}\right\}
\end{aligned}$$

Teorema 1.2.5 (De Unicidad.) *Una función de distribución está determinada de manera única por su función característica e inversamente, mediante la fórmula de inversión, (caso discreto).*

Fórmula de Inversión

Si X es una v.a. discreta que toma valores enteros con función característica dada por

$$\varphi_X(t) = \sum_{-\infty}^{\infty} e^{itj} f_X(j),$$

entonces la densidad es

$$f_X(k) = P(x = k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\varphi_X(t)dt.$$

Demostración.

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\varphi_X(t) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\left(\sum_{j\in\mathbb{Z}} e^{itj}P(X=j)\right)dt \qquad (1.2.11)\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{j} e^{i(j-k)t}P(X=j)dt\\
&= \frac{1}{2\pi}\sum_{j\in\mathbb{Z}} P(X=j)\int_{-\pi}^{\pi} e^{i(j-k)t}dt\\
&= \sum_{j} P(X=j)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(j-k)t}dt
\end{aligned}$$

pero

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(j-k)t}dt = \begin{cases} 0 & \text{si } j \neq k \\ 1 & \text{si } j = k \end{cases}$$

Si $j \neq k$ entonces la integral es igual a

$$\begin{aligned}
\left[\frac{1}{2}\pi i(j-k)e^{i(j-k)t}\right]_{-\pi}^{\pi} &= \frac{1}{2}\pi i(j-k)\left[\exp\{i(j-k)\pi\} - exp\{-i(j-k)\pi\}\right]\\
&= \frac{\sin[(j-k)\pi]}{\pi(j-k)} = 0,
\end{aligned}$$

dado que $\sin(m\pi) = 0$ para todo m entero.

Si $j = k$, entonces

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-k)t}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(0)}dt$$
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} dt = 1,$$

lo que implica que (1.2.12) es igual a

$$\sum_{j\in\mathbb{Z}} P(X = j) I_{\{k\}}(j) = P(X = k)\,.$$

Teorema 1.2.8 *Sea X una v.a. cuya función característica es integrable, esto es*

$$\int_{-\infty}^{\infty} |\varphi_X(t)|dt < \infty$$

se puede demostrar que si X es v.a. continua, la densidad $f_X(x)$ estará dada por

$$f_X(x) = \frac{1}{2\pi}\int_{\infty}^{-\infty} e^{-itX}\varphi_X(t)dt.$$

La demostración de esta fórmula de inversión no se incluye por no estar al alcance de este trabajo.

Ejemplo 1.2.4

Sea $X \sim N(0,\sigma^2)$ se utilizará la fórmula anterior para obtener la densidad. Se sabe que $\varphi_X(t) = \exp\{-\sigma^2t^2/2\}$, entonces por definición de función característica se tiene

$$e^{-\sigma^2t^2/2} = \int_{-\infty}^{\infty} e^{itx}\frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}dx.$$

Si se reemplaza t por $-t$ y σ por $\frac{1}{\sigma}$ en esta fórmula, entonces

$$e^{-t^2/2\sigma^2} = \int_{-\infty}^{\infty} e^{-itx}\frac{\sigma}{\sqrt{2\pi}}e^{-\sigma^2x^2/2}dx$$

o equivalentemente

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-t^2/2\sigma^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-\sigma^2x^2/2}dx$$

finalmente si se intercambia el papel de los símbolos x y t en la última ecuación, se obtiene

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{itx}e^{-\sigma^2t^2/2}dt$$

la cual es justamente la fórmula de inversión para el caso continuo, y entonces

$$f_X(x) = \frac{1}{\sigma\sqrt{s\pi}}e^{-x^2/2\sigma^2}.$$

Teorema 1.2.7 *Si dos v.a. tienen la misma función característica, entonces ambas tienen la misma función de distribución.*

Ejemplo 1.2.5

Sean X y Y v.a.i. con distribución $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$ respectivamente, entonces la distribución de $Z = X + Y$ es

$$\begin{aligned}
\varphi_X(t) &= \exp\{i\mu_1 t\}\exp\{-\sigma_1^2t^2/2\} \\
\varphi_Y(t) &= \exp\{i\mu_2 t\}\exp\{-\sigma_2^2t^2/2\} \\
&\Rightarrow \varphi_{X+Y} = \varphi_X(t)\varphi_Y(t) \\
&= \exp\{i(\mu_1+\mu_2)t\}\exp\left\{-\frac{-(\sigma_1^2+\sigma_2^2)t^2}{2}\right\} \\
&= \exp\{i(\mu^*)t - \frac{\sigma_*^2t^2}{2}\} = \varphi_Z(t)
\end{aligned}$$

donde $Z \sim N(\mu^*, \sigma_*^2)$ y $\mu^* = \mu_1 + \mu_2$ y $\sigma_*^2 = \sigma_1^2 + \sigma_2^2$

Teorema 1.2.8 (De Continuidad de Levy.) *Una sucesión de funciones de distribución $\{F_n(\cdot)\}$ tiene distribución límite F, es decir, converge a $F(\cdot)$ en todos los puntos de continuidad de $F(\cdot)$ si y solo si, las correspondientes funciones características $\varphi_n(t)$ converge a $\varphi(t)$, donde $\varphi(t)$ es una función continua en $t = 0$. En este caso $\varphi(\cdot)$ es la función característica de $F(\cdot)$.*

Ejemplo 1.2.6

Si $X_n \sim Bin(n, p)$, entonces la sucesión

$$Y_n = \frac{X_n - np}{\sqrt{np(1-p)}}, \quad 0 < p < 1$$

tiene como límite a la distribución $N(0,1)$ cuando $n \to \infty$

$$\varphi_{X_n}(t) = (pe^{it} + q)^n, \quad \text{con } p + q = 1$$

por lo tanto la función característica $\varphi_{Y_n}(t)$ de Y_n está dada por

$$\begin{aligned}
\varphi_{Y_n}(t) &= \left(p\exp\left\{\frac{it}{\sqrt{npq}}\right\} + q\right)^n \exp\left\{\frac{-it}{npq}\right\} \\
&\Rightarrow \varphi_{Y_n}(t)\varphi_{X_n}\left(\frac{t}{\sqrt{npq}}\right)\exp\left\{\frac{-itnp}{\sqrt{npq}}\right\} \\
&= \left(p\exp\left\{\frac{it}{\sqrt{npq}}\right\}\exp\left\{\frac{-itp}{\sqrt{npq}}\right\} + q\exp\left\{\frac{-itp}{\sqrt{npq}}\right\}\right)^n \\
&= \left(p\exp\left\{\frac{itq}{\sqrt{npq}}\right\} + q\exp\left\{\frac{-itp}{\sqrt{npq}}\right\}\right)^n \\
&= \left(\underbrace{p\exp\left\{it\sqrt{\frac{q}{np}}\right\} + q\exp\left\{-it\sqrt{\frac{p}{nq}}\right\}}_{f(t)}\right)^n
\end{aligned}$$

ahora desarrollando las derivadas de $f(t)$ y evaluando en 0 se tiene que

$$\begin{aligned}
f(0) &= p + q = 1 \\
f'(0) &= pi\sqrt{\frac{q}{np}} - qi\sqrt{\frac{p}{nq}} = 0 \\
f''(0) &= \frac{-p}{n} - \frac{q}{n} = -\frac{1}{n}
\end{aligned}$$

por lo tanto

$$f(t) = 1 + 0 \cdot t - \frac{1}{n}t^2 + R_n$$

lo que implica que

$$\left(q \exp\left\{-it\sqrt{\frac{p}{nq}}\right\} + p \exp\left\{it\sqrt{\frac{q}{np}}\right\}\right)^n = \left(1 - \frac{t^2}{2n} + R_n\right)^n \to e^{-\frac{t^2}{2}} \tag{1.2.12}$$

cuando $n \to \infty$, y esto debido a que

$$\lim_{n\to\infty}\left(1 + \frac{a}{n}\right)^n = e^a,$$

donde $\exp\{-\frac{t^2}{2}\}$ es la función característica de una distribución $N(0,1)$, y de aquí que $Y_n \xrightarrow{d} N(0,1)$ por el Teorema de Continuidad.

1.3 Desigualdades Útiles

Teorema 1.3.1 (Desigualdad de Tchebyshev) *Sea X una v.a. y $g(\cdot)$ una función no negativa con dominio la recta real; entonces*

$$P(g(X) \geq k) \leq \frac{E(g(x))}{k}$$

para toda $k > 0$.

Demostración Suponga que X es continua con densidad de probabilidad $f_X(\cdot)$, entonces

$$\begin{aligned} E(g(X)) &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ &= \int_{\{x:g(x)\geq k\}} g(x) f_X(x) dx + \int_{\{x:g(x)<k\}} g(x) f_X(x) dx \\ &\geq \int_{\{x:g(x)\geq k\}} g(x) f_X(x) dx \\ &\geq \int_{\{x:g(x)\geq k\}} k f_X(x) dx \\ &= kP(g(X)\geq k) \end{aligned}$$

de manera que

$$P(g(X)\geq k) \leq \frac{E(g(X))}{k}.$$

□

Una prueba similar se sigue para el caso discreto.

Corolario 1.3.1 (De la Desigualdad de Tchebyshev) *Si X es una v.a. con varianza finita*

$$P[|X-\mu_X| \geq r\sigma_X] = P[(X-\mu_X)^2 \geq r^2\sigma_X^2] \leq \frac{1}{r^2}. \tag{1.3.13}$$

para cada $r > 0$

Demostración.

Si se toma $g(x) = (x-\mu_X)^2$ y $k = r^2\sigma_X^2$ el resultado se sigue del teorema anterior.

Observación:

Si X es una v.a. con varianza finita

$$P[|X-\mu_X| < r\sigma_X] \geq 1 - \frac{1}{r^2}$$

que se obtienen solo reescribiendo (1.3.13).

Esta desigualdad es útil ya que en la Capítulo 3 se utilizará para porbar las leyes de los grandes números. La desigualdad proporciona cotas que no dependen de la distribución de la v.a. X para las probabilidades de eventos particulares que se describen en términos de una variable aleatoria. la media y la varianza.

Teorema 1.3.2 (Desigualdad de Kolmogorov) *Sean $X_1, X_2, \ldots, X_n$ v.a.i. con medias $\mu_1, \mu_2, \ldots, \mu_n$ y varianza $\sigma_1^2, \ldots, \sigma_n^2$. Sean $S_k = X_1 + X_2 + \cdots + X_k$, $k = 1, \ldots, n$ y considere la v.a.*

$$\max_{k=1,\ldots,n} \left\{ S_k - \sum_{i=1}^{k} \mu_i \right\},$$

que da la desviación máxima de las v.a. S_k, de sus correspondientes medidas, entonces

$$P\left\{ \max_{k=1,\ldots,n} \left| S_k - \sum_{i=1}^{k} \mu_i \right| < a\left(\sigma_1^2 + \cdots \sigma_n^2\right)^{\frac{1}{2}} \right\} \geq 1 - \frac{1}{a^2}$$

para cada número a positivo.

Demostración.

Obsérvese que el evento

$$\left\{ \max_{k=1,\ldots,n} \left| S_k - \sum_{i=1}^{k} \mu_i \right| < a\left(\sigma_1^2 + \cdots + \sigma_n^2\right)^{\frac{1}{2}} \right\}$$

ocurre si y sólo si

$$\left| S_k - \sum_{i=1}^{k} \mu_i \right| < a\left(\sigma_1^2 + \cdots + \sigma_n^2\right)^{\frac{1}{2}} \tag{1.3.14}$$

para toda $k = 1, \ldots, n$. Este evento puede descomponerse de la siguiente manera. Sea

$Z_k(w) = 1$, si w es tal que

$$\left|S_r(w) - \sum_{i=1}^{r} \mu_i\right| < a(\sigma_1^2 + \cdots + \sigma_n^2)^{\frac{1}{2}} \text{ para toda } r \leq k-1$$

$$\left|S_k - \sum_{i=1}^{k} \mu_i\right| \geq a(\sigma_1^2 + \cdots + \sigma_n^2)^{\frac{1}{2}}$$

y $Z_k(w) = 0$ en otro caso. Así $Z_k(w)$ es la v.a. indicadora del conjunto de puntos w para los que las desigualdades (1.3.14) se violan por la k-ésima desigualdad pero no por un índice menor k.

Para cualquier w, o todas las Z_k son cero o exactamente una de ellas es igual a 1, y el evento

$$\left\{\max_{k=1,\ldots,n} \left|S_k - \sum_{i=1}^{k} \mu_i\right| < a\left(\sigma_1^2 + \cdots + \sigma_n^2\right)^{\frac{1}{2}}\right\} = \{Z_1 + \cdots + Z_n = 0\}.$$

Se probará equivalentemente, que el complemento tiene probabilidad a lo más igual a $\frac{1}{a^2}$.

$$\{Z_1 + \cdots + Z_n = 0\}^c = \{Z_1 + \cdots + Z_n = 1\}$$

Obsérvese que

$$E(Z_1 + \cdots + Z_n) = P(Z_1 + \cdots + Z_n = 1),$$

ya que la suma vale 0 ó 1.

Por lo que se demostrará que

$$E(Z_1 + \cdots + Z_n) \leq \frac{1}{a^2}$$

Obsérvese que

$$\begin{aligned}
\sum_{k=1}^{n} E\left[Z_k\left(S_n - \sum_{i=1}^{n}\mu_i\right)^2\right] &= E\left[\sum_{k=1}^{n} Z_k\left(S_n - \sum_{i=1}^{n}\mu_i\right)^2\right] \\
&\leq E\left[\sum_{k=1}^{n}\left(S_n - \sum_{i=1}^{n}\mu_i\right)^2\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} E\left[S_n - \sum_{i=1}^{n}\mu_i\right]^2 \\
&= \sum_{k=1}^{n}\sigma_k^2 \qquad (1.3.15)
\end{aligned}
$$

ya que a lo más una de las $Z_k = 1$.

Los términos individuales

$$Z_k\left(S_n - \sum_{i=1}^{n}\mu_i\right)^2$$

pueden ser escritos como

$$
\begin{aligned}
Z_k\left(S_n - \sum_{i=1}^{n}\mu_i\right)^2 &= Z_k\left\{\left[\left(\sum_{i=1}^{k}x_i - \sum_{i=1}^{k}\mu_i\right) + \left(\sum_{j=k+1}^{n} x_j - \sum_{j=k+1}^{n}\mu_j\right)\right]^2\right\} \\
&= Z_k\left(S_k - \sum_{i=1}^{k}\mu_i\right)^2 + \\
&+ 2Z_k\left(S_k - \sum_{i=1}^{k}\mu_i\right)\left(\sum_{j=k+1}^{n} X_j - \sum_{j=k+1}^{n}\mu_j\right) \\
&+ Z_k\left(\sum_{j=k+1}^{n} X_j - \sum_{j=k+1}^{n}\mu_j\right)^2,
\end{aligned}
$$

pero Z_k es una v.a. definida en términos de $X_1, \ldots, X_k$ solamente, por eso es independiente de $X_{k+1}, \ldots, X_n$. Por lo tanto

$$
\begin{aligned}
&E\left[Z_k\left(S_k - \sum_{i=1}^{k}\mu_i\right)\left(\sum_{j=k+1}^{n} X_j - \sum_{j=k+1}^{n}\mu_j\right)\right] \\
&= E\left[Z_k\left(S_k - \sum_{i=1}^{k}\mu_i\right)\right]\underbrace{E\left[\sum_{j=k+1}^{n} X_j - \sum_{j=k+1}^{n}\mu_j\right]}_{cero} = 0
\end{aligned}
$$

Usando nuevamente la independencia se tiene que

$$E\left(Z_k\left(\sum_{j=k+1}^{n}X_j-\sum_{j=k+1}^{n}\mu_j\right)^2\right) = E(Z_k)E\left(\sum_{j=k+1}^{n}X_j-\sum_{j=k+1}^{n}\mu_j\right)^2$$
$$= E(Z_k)\left(\sigma_{k+1}^2+\cdots+\sigma_n^2\right)\geq 0$$

por lo tanto

$$E\left[Z_k\left(S_n-\sum_{i=1}^{n}\mu_i\right)^2\right]\geq E\left[Z_k\left(S_k-\sum_{i=1}^{k}\mu_i\right)^2\right].$$

La esperanza del lado derecho puede escribirse,

$$E\left[Z_k\left(S_k-\sum_{i=1}^{k}\mu_i\right)^2\right] = E\left[Z_kI_{(Z_k=1)}(S_k-\mu_1-\cdots-\mu_k)^2\right] + \underbrace{E\left[Z_kI_{(Z_k=0)}(S_k-\mu_1-\cdots-\mu_k)^2\right]}_{Z_k=0}.$$

Sobre el conjunto $Z_k=1$ se tiene por definición que

$$|S_k-\mu_1-\cdots-\mu_k|>a(\sigma_1^2+\cdots+\sigma_n^2)^{1/2}$$

por lo que

$$Z_kI_{(Z_k=1)}\left(S_k-\sum_{i=1}^{k}\mu_i\right)^2>a^2(\sigma_1^2+\cdots+\sigma_n^2)Z_k.$$

Tomando las esperanzas de ambos lados

$$E\left[Z_k(S_k-\mu_1-\cdots-\mu_k)^2\right]\geq a^2(\sigma_1^2+\cdots+\sigma_n^2)E(Z_k),$$

y por lo tanto

$$\sum_{i=1}^{n}E\left[Z_k(S_k-\mu_1-\cdots-\mu_k)^2\right]\geq a^2(\sigma_1^2+\cdots+\sigma_n^2)\sum_{i=1}^{n}E(Z_k),$$

utilizando (1.3.15) se sabe que

$$\sum_{i=1}^{n}E\left(Z_k(S_k-\mu_1-\cdots-\mu_k)^2\right)\leq\sum_{i=1}^{n}\sigma_i^2$$

de aquí que

$$\sigma_1^2 + \cdots + \sigma_n^2 \geq a^2(\sigma_1^2 + \cdots + \sigma_n^2)E(Z_1 + \cdots + Z_n)$$

lo que implica que

$$E(Z_1 + \cdots + Z_n) \leq \frac{1}{a^2},$$

y como

$$P(Z_1 + \cdots + Z_n = 1) = E(Z_1 + \ldots + Z_n),$$

entonces

$$P(Z_1 + \cdots + Z_n = 1) = \frac{1}{a^2},$$

de aquí que

$$\begin{aligned}
P\left(\max_{k=1,\ldots,n}\left|S_k - \sum_{i=1}^{k}\right| < a(\sigma_1^2 + \cdots + \sigma_k^2)^{1/2}\right) &= P(Z_1 + \cdots + Z_n = 0) \\
&= 1 - P\left(\sum_{i=1}^{n} Z_i = 1\right) \\
&= 1 - \frac{1}{a^2}
\end{aligned}$$

y entonces

$$P\left(\max_{k=1,\ldots,n}\left|S_k - \sum_{i=1}^{k}\mu_i\right| < a\left(\sigma_1^2 + \cdots + \sigma_k^2\right)^{1/2}\right) = 1 - \frac{1}{a^2}$$

□

Observaciones. Para $n = 1$, la desigualdad de *Kolmogoroff* se reduce a la desigualdad de *Tchebyshev*.

1.4 Lemas Útiles.

Otro resultado que será útil en el Capítulo 3 es el conocido *Lema de Borel-Cantelli*; antes de enunciarlo se introducirá la siguiente notación.

Si $\{A_n : n \in \mathbf{N}\}$ es una sucesión de eventos, se sabe que

$$\limsup\{A_n\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Si $w \in \limsup\{A_n\}$ esto implica que $w \in \bigcup_{k=n}^{\infty} A_k$ para toda $n \in \mathbf{N}$, lo que sinifica que w está en un número infinito de A_k's. Así la siguiente notación tiene sentido:

$$\limsup\{A_n\} = \{w \in \Omega : w \in A_n \ i.v.\},$$

donde $i.v.$ quiere decir infinitas veces. El siguiente resultado que involucra esta idea es fundamental en la Teoría de Probabilidad.

Lema 1.4.1 (Borel-Cantelli) *Si $\{A_n : n \in \mathbf{N}\}$ es tal que $\sum_{n=1}^{\infty} P(A_n) < \infty$ entonces $P(A_n : i.v.) = 0$. Si $\{A_n : n \in \mathbf{N}\}$ son independientes y $\sum_{n=1}^{\infty} P(A_n) = \infty$ entonces $P(A_n : \mathrm{i.v.}) = 1$*

Demostración.

Si $A = \{A_n : \mathrm{i.v.}\}$ entonces $A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$. Entonces $A \subset \bigcup_{k=n}^{\infty} A_k$ para toda $n \in \mathbf{N}$. De aquí que

$$P(A) \leq \sum_{k=n}^{\infty} P(A_k) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k),$$

si $\sum_{n=1}^{\infty} P(A_n) < \infty$ entonces $\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$ y por lo tanto $P(A) = 0$

Por otro lado $A^c \subset \bigcap_{k=n}^{\infty} A_k^c$ y si $\{A_n : n \in \mathbf{N}\}$ son independientes, entonces

$$P(A^c) \leq \prod_{k=n}^{\infty} (1 - P(A_k)) \leq \exp\left\{-\sum_{k=n}^{\infty} P(A_n)\right\},$$

pues $(1 - x) \leq e^{-x}$ para toda $x \in \mathbf{R}$, y si $\sum_{n=1}^{\infty} P(A_n) = \infty$ se sigue que $P(A) = 1$

□

Un resultado que proporciona otra caracterización de convergencia casi segura y que será necesario para la demostración de la *Ley Fuerte de los Grande Números* en el caso Binomial es el siguiente,

Lema 1.4.2 *Si* $\{X_n : n \in \mathbf{N}\}$ *y* X *son v.a., entonces*

$$P\left[\lim_{n\to\infty} X_n = X\right] = 1 \textit{ si y sólo si } P\left[|X_n - X| > \frac{1}{k} i.v.\right] = 0 \textit{ para toda } k \in \mathbf{N}$$

Demostración.

Sea $A_k = \{w \in \Omega : |X_n(w) - X(w)| > \frac{1}{k} i.v.\}$ para toda $k \in \mathbf{N}$ y hágase $A = \bigcup_{k=1}^{\infty} A_k$.

Si $w \notin A$ se sigue que $|X_n(w) - X(w)| > \frac{1}{k}$ para toda $k \in \mathbf{N}$ se cumple para un númer finito de índices y entonces $\lim_{n\to\infty} X_n = X$. Recíprocamente, si $\lim_{n\to\infty} X_n = X$ entonces $w \notin A_k$ para toda $k \in \mathbf{N}$ por lo que $w \notin A$ y por lo tanto $A^c = \{w \in \Omega : \lim_{n\to\infty} X_n(w) = X(w)\}$. Entonces

$$P(\lim_{n\to\infty} X_n = X) = 1 \text{ si y sólo si } P(A) = 0$$

y como $A_k \subset A$ para toda $k \in \mathbf{N}$ se tiene que $P(A) = 0$ para toda $k \in \mathbf{N}$.

□

Capítulo 2

Conceptos básicos y aplicaciones

Las ideas que se presentan en este capítulo tienen como objetivo aclarar algunos de los conceptos básicos utilizados en las aplicaciones del ambiente asegurador, siendo base fundamental para la obtención de futuros resultados que constituirán la aportación principal de este trabajo.

2.1 Sumas de variables aleatorias.

Existe en la Teoría de Probabilidad y Estadística, un problema interesante, que consiste en calcular la distribución de sumas de variables aleatorias independientes y con la misma distribución debido a que dicha distribución es la base de importantes resultados *límite* comunmente usados en la práctica. Supóngase que X y Y son variables aleatorias positivas independientes, continuas, y no necesariamente con la misma distribución con funciones de densidades $f_X(x)$ y $f_Y(y)$ respectivamente.

Defínase $Z = X + Y$, como la suma de variables aleatorias, entonces la función de distribución de Z está dada por:

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X + Y \leq z) \\ &= \sum_{y \leq z} P(X + y \leq z | Y = y) P(Y = y) \end{aligned}$$

$$= \sum_{y \leq z} P(X \leq z - y | Y = y) P(Y = y),$$

por independencia de X y Y

$$F_Z(z) = \sum_{y \leq z} F_X(z - y) f_Y(y)$$

y

$$f_Z(z) = \sum_{y \leq z} f_X(z - y) f_Y(y).$$

Para el caso continuo se tiene de igual forma

$$\begin{aligned} F_Z(z) &= \int_0^z P(X \leq z - y \mid Y = y) f_Y(y) dy \\ &= \int_0^z F_x(z - y) f_y(y) dy \end{aligned}$$

y

$$f_Z(z) = \int_0^z f_X(z - y) f_y(y) dy.$$

En resumen, para v.a. no negativas,

$$F_Z(z) = \begin{cases} \sum_{y \leq z} F_X(z - y) f_Y(y) & \text{para el caso discreto} \\ \int_0^z F_X(z - y) f_y(y) dy & \text{para el caso continuo} \end{cases} \tag{2.1.1}$$

y

$$f_Z(z) = \begin{cases} \sum_{y \leq z} f_X(z - y) f_Y(y) & \text{para el caso discreto} \\ \int_0^z f_X(z - y) f_y(y) dy & \text{para el caso continuo} \end{cases} \tag{2.1.2}$$

Ahora bien, de la teoría del cálculo se sabe que dadas dos fuciones f y g unidimensionales, la función

$$h(z) = \int_{-\infty}^{+\infty} f_X(x) g(z - x) dx,$$

con $-\infty < z < +\infty$ y f_X y f_Y funciones de densidad integrables, es conocida como la convolución de f y g o la convolución de orden 1 de g, denotada por $h = g * f$, y leída como *g convolucionada con f*. De (2.1.2), se observa entonces que la función de densidad de la suma de dos v.a. $Z = X + Y$ es la convolución de las densidades f_X y f_Y.

Ejemplo 2.1.1

Si X y Y son variables aleatorias independientes con distribución gamma con parámetros (s, λ) y (t, λ) respectivamente, entonces $Z = X + Y$ es una variable aleatoria gamma con parámetros $(s + t, \lambda)$.

Demostración.

$$
\begin{aligned}
f_Z(a) &= \int_0^a f_X(a-y) f_Y(y)\,dy \\
&= \int_0^a \frac{1}{\Gamma(s)} \lambda^s e^{-\lambda(a-y)}(a-y)^{s-1} \frac{1}{\Gamma(t)} \lambda^t y^{t-1} e^{-\lambda y}\,dy \\
&= \int_0^a \frac{1}{\Gamma(s)\Gamma(t)} \lambda^s \lambda^t e^{-\lambda a}(a-y)^{s-1} y^{t-1}\,dy \\
&= k e^{-\lambda a} a^{s+t-1} \int_0^a \left(\frac{a-y}{a}\right)^{s-1} \left(\frac{y}{a}\right)^{t-1} dy \\
&= c e^{-\lambda a} a^{s+t-1} \int_0^1 (1-z)^{s-1} z^{t-1}\,dz \\
&= c e^{-\lambda a} a^{s+t-1}, \text{ con } k = \text{cte.}
\end{aligned}
$$

ya que $(1-z)^{s-1} z^{t-1}$ es una densidad Beta integrada sobre todo su recorrido.

Debe notarse que $e^{-\lambda a} a^{s+t-1}$ es el núcleo de una densidad gamma con parámetro $(s + t, \lambda)$, entonces

$$f_Z(a) = c e^{-\lambda a} a^{s+t-1},$$

donde k debe ser tal que $f_Z(a)$ integrada sobre todo su recorrido sea igual a uno, por lo tanto

$$f_Z(a) = \frac{1}{\Gamma(s+t)} e^{-\lambda a} (\lambda a)^{s+t-1}.$$

La suma de variables aleatorias independientes, y más específicamente, con la misma distribución, puede ser extendida a más de dos variables.

Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de una variable aleatoria Y. A menos que se haga explicito en otro caso, se supondrá que los momentos de todos los órdenes existen y en particular, $E(Y) = \mu$ y $Var(Y) = \sigma^2 < \infty$, definiendo

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

entonces

$$E(S_n) = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mu = n\mu \tag{2.1.3}$$

y

$$Var(S_n) = Var\left(\sum_{i=1}^{n} Y_i\right) = \sum Var(Y_i) = n\sigma^2. \tag{2.1.4}$$

La densidad exacta de S_n ($f_{S_n}(s)$), extendiendo el resultado anterior, se calcula como

$$f_{S_n}(s) = \int f_Y(y_1) \cdots f_Y(y_{n-1}) f_Y(s - y_1 - y_2 - \cdots - y_{n-1}) dy_1 \cdots dy_{n-1}. \tag{2.1.5}$$

con una correspondiente suma en el caso discreto.

Obsérvese que la distribución de la suma S_n, es la convolución de orden n de la densidad $f_Y(y)$. Es claro por (2.1.5) que a medida que n crece, es analíticamente más difícil el cálculo de la distribución debido a la complejidad de resolver convoluciones de orden mayor que uno, pues el factor integración está incluido en este proceso. Para determinar la distribución de las suma de más de dos v.a.i. se puede usar el proceso de convolución iterativo, de manera relativamente fácil si las densidades de las v.a. en juego son discretas. Así para $S = X_1 + X_2 + \cdots + X_n$ donde las v.a.

X_i's son independientes, con distribución F_i, donde $F^{(k)}$ es la función de la distribución de

$$X_1 + X_2 + \cdots + X_k,$$

se tendría aplicando (2.1.1) de manera secuencial

$$\begin{aligned}
F^{(1)} &= F_1 \\
F^{(2)} &= F_2 * F^{(1)} = F_2 * F_1 \\
F^{(3)} &= F_3 * F^{(2)} \\
F^{(4)} &= F_4 * F^{(3)} \\
&\vdots \\
F_S &= F^{(n)} = F_n * F^{(n-1)}.
\end{aligned}$$

Ejemplo 2.1.2

Las v.a. X_1, X_2 y X_3 son independientes con densidad definida en las columnas (1), (2) y (3) de la Tabla (2.1.1), lo que se desea encontrar es la distribución y la densidad de $S = X_1 + X_2 + X_3$.

Solución.

1. **Columnas (1) - (3) son la información proporcionada**
2. **Columna (4) es la distribución acumulada de la columna (1)**
3. **Columna (5) se deriva de la columna (2) y (4) usando (2.1.1)**
4. **Columna (6) se deriva de la columna (3) y (5) usando (2.1.1)**

El cálculo de la columna (6) completa la obtención de la distribución de S y la función de densidad se obtiene por diferencias. Como ilustración se incluyen las columnas (7) y (8), utilizando las fórmulas para el cálculo de densidades dadas en (2.1.2).

Tabla 2.1.1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
X	f^1	f^2	f^3	F^1	F^2	F^3	f^2	f^3
0	0.4	0.5	0.6	0.4	0.20	0.120	0.20	0.120
1	0.3	0.2	0.0	0.7	0.43	0.258	0.23	0.138
2	0.2	0.1	0.1	0.9	0.63	0.398	0.20	0.140
3	0.1	0.1	0.1	1.0	0.79	0.537	0.16	0.139
4		0.1	0.1	1.0	0.90	0.666	0.11	0.129
5			0.1	1.0	0.96	0.781	0.06	0.115
6				1.0	0.99	0.896	0.03	0.088
7				1.0	1.00	0.928	0.01	0.053
8				1.0	1.00	0.964	0.00	0.036
9				1.0	1.00	0.985	0.00	0.020
10				1.0	1.00	0.995	0.00	0.009
11				1.0	1.00	0.999	0.00	0.003
12				1.0	1.00	1.000	0.00	0.000

$$\begin{aligned} F^{(2)}(0) &= \sum_{y\leq 1} F_X(0-y)f_Y(y) \\ &= F_X(0)f_y(0) = (0.4)(0.5) \\ &= 0.20 \end{aligned}$$

$$\begin{aligned} F^{(2)}(1) &= \sum_{y\leq 1} F_X(1-y)f_Y(y) \\ &= F_X(1)f_y(0) + F_X(0)f_Y(1) \\ &= 0.43 \end{aligned}$$

$$\begin{aligned} F^{(2)}(2) &= F_X(2)f_y(0) + F_X(1)f_Y(1) + F_X(0)f_Y(2) \\ &= 0.45 + .14 + 0.04 \\ &= 0.63 \end{aligned}$$

$$\begin{aligned} F^{(2)}(3) &= F_X(3)f_y(0) + F_X(2)f_Y(1) + F_X(1)f_Y(2) + F_X(0)f_Y(3) \\ &= 0.5 + 0.18 + 0.07 + 0.04 \\ &= 0.79 \end{aligned}$$

$$\begin{aligned} F^{(2)}(4) &= F_X(4)f_y(0) + F_X(3)f_Y(1) + F_X(2)f_Y(2) \\ &+ F_X(1)f_Y(3) + F_X(0)f_Y(4) \\ &= 0.5 + 0.2 + 0.09 + 0.07 + 0.04 \\ &= 0.90 \end{aligned}$$

$$\begin{aligned} F^{(2)}(5) &= F_X(5)f_y(0) + F_X(4)f_Y(1) + F_X(3)f_Y(2) \\ &+ F_X(2)f_Y(3) + F_X(1)f_Y(4) + F_X(0)f_Y(5) \\ &= 0.5 + 0.2 + 0.1 + 0.09 + 0.07 \\ &= 0.96 \end{aligned}$$

$$\begin{aligned} F^{(2)}(6) &= F_X(6)f_y(0) + F_X(5)f_Y(1) + F_X(4)f_Y(2) \\ &+ F_X(3)f_Y(3) + F_X(2)f_Y(4) + F_X(1)f_Y(5) \\ &+ F_X(0)f_Y(6) \\ &= 0.5 + 0.2 + 0.1 + 0.1 + 0.09 \\ &= 0.99 \end{aligned}$$

$$\begin{aligned} F^{(2)}(7) &= F_X(7)f_y(0) + F_X(6)f_Y(1) + F_X(5)f_Y(2) \\ &+ F_X(4)f_Y(3) + F_X(3)f_Y(4) + F_X(2)f_Y(5) \\ &+ F_X(1)f_Y(6) + F_X(0)f_Y(7) \\ &= 0.5 + 0.2 + 0.1 + 0.1 + 0.1 \\ &= 1.0 \end{aligned}$$

Obsérvese que para $x > 7$, se tiene que $F^{(2)}(x) = 1.0$

Para resolver (2.1.5), en los casos en que la distribución de las X_i's no es simple (como en el caso de una distribución discreta) se recurre a métodos numéricos, los que no están incluidos en este trabajo, que se encargan de encontrar una densidad para la cual se puedan calcular las convoluciones en forma simple, y entonces escribir,

$$f_Y(y) = f_0(y) + \delta(y),$$

donde $\delta(y)$ es una perturbación pequeña, o bien hacer

$$f_Y(y) = f_0(y)\left[1 + \sum c_r p_r(y)\right],$$

donde $\{c_r p_r(y)\}$ es un conjunto de polinomios ortogonales, que pueden ser, por ejemplo, los *Polinomios de Hermite-Chebyshev*, que serán descritos en el próximo capítulo, asociados con $f_0(y)$. donde esta densidad puede ser *Normal, Gamma, Poisson o Binomial.*

Una consecuencia de la complejidada de la distribución anterior, es la idea de enfocar el problema encontrando la distribución de S_n en el límite, es decir, abordar el cálculo de la distribuición asintótica de S_n. Este problema no es sencillo, y dependerá probablemente de ciertas propiedades de las variables aleatorias involucradas. Este estudio se realizará en el capítulo 3.

Antes de continuar, es conveniente marcar que el primer paso para cualquier cálculo asintótico es estandarizar la variable aleatoria de interés, permitiendo de esta manera cambiar tanto de origen como de escala, obteniendo así una variable aleatoria con media igual a cero y varianza igual a uno. En este caso la variable aleatoria es S_n la cual al ser estandarizada se expresa de la siguiente forma,

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \tag{2.1.6}$$

donde, $E(S_n^*) = 0$ y $Var(S_n^*) = 1$.

Esta estandarización es pieza fundamental en el desarrollo del siguiente capítulo. Cabe mencionar que cuando se utilizan variables aleatorias cuyos dos primeros momentos no existen, será necesario un tratamiento alternativo para su estudio.

2.2 Momentos, cumulantes y sus funciones generadoras de variables aleatorias.

Los momentos y cumulantes son un conjunto de constantes que describen de cierta manera la distribución de una variable aleatoria. Algunos investigadores los utilizan para ir marcando de manera más precisa la forma de alguna distribución. Es importante mencionar que sólo un grupo reducido de momentos y cumulantes tienen una interpretación física, la cual se explicará posteriormente.

2.2.1 Momentos y su función generadora.

Continuando con el caso de $Y_1, Y_2, \ldots, Y_n$ que es una muestra aleatoria de una de una variable aleatoria Y; se definirá a la función generadora de momentos de esa variable aleatoria Y como

$$M(Y;t) = M(t) = E(e^{tY}), \tag{2.2.7}$$

denotando al r-ésimo momento de Y alrededor del cero como,

$$\mu'_r = E(Y^r), \tag{2.2.8}$$

y al r-ésimo momento de Y alrededor de la media como,

$$\mu_r = E[(Y-\mu)^r]. \tag{2.2.9}$$

Para desarrollar la expresión (2.2.7) se supondrá que $M(t)$ es convergente en un intervalo de valores reales de t que contienen al punto $t = 0$ en su interior, de esta forma desarrollando por la expansión en serie de Taylor alrededor del cero se tiene que

$$\begin{aligned} M(Y;t) = E(e^{tY}) &= E\left(\sum_{j=0}^{\infty}\frac{(tY)^j}{j!}\right) \\ &= 1 + \frac{t}{1!}E(Y) + \frac{t^2}{2!}E(Y^2) + \cdots \\ &\overset{por(\ 2.2.8)}{=} 1 + \frac{t}{1!}\mu'_1 + \frac{t^2}{2!}\mu'_2 + \cdots \\ &= 1 + \sum_{r=1}^{\infty}\frac{t^r}{r!}\mu'_r. \end{aligned} \tag{2.2.10}$$

Si se resta el escalar μ a la v.a. Y, la expresión anterior queda de la siguiente forma,

$$
\begin{aligned}
M(Y-\mu;t) &= E(e^{t(Y-\mu)}),\\
&= E\left(\sum_{j=0}^{\infty}\frac{(t(Y-\mu))^j}{j!}\right)\\
&= 1+E\left(\frac{t(Y-\mu)}{1!}+\frac{t^2(Y-\mu)^2}{2!}+\cdots\right)\\
&\overset{por(2.2.9)}{=} 1+\frac{t}{1!}+\frac{t^2}{2!}\mu_2+\frac{t^3}{3!}\mu_3+\cdots\\
&= 1+\sum_{r=1}^{\infty}\frac{t^r}{r!}\mu_r.
\end{aligned}
$$

Ejemplo 2.2.1 (Distribución Normal)

Sea Y una v.a. con distribución normal con media μ y varianza σ^2, entonces

$$
\begin{aligned}
M(Y;t) &= E\left(e^{tY}\right)=E\left(e^{t(Y-\mu)}\right)e^{t\mu}\\
&= \int_{\infty}^{-\infty}\exp\{t(y-\mu)\}\exp\{t\mu\}\, f_Y(y)dy\\
&= \frac{e^{t\mu}}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{\frac{-1}{2\sigma^2}\left[(Y-\mu)^2-2\sigma^2t(Y-\mu)\right]\right\}dy,
\end{aligned}
$$

si se completa el cuadrado en los corchetes, se tiene,

$$
\begin{aligned}
(Y-\mu)^2-2\sigma^2t(Y-\mu) &= (Y-\mu)^2-2\sigma^2t(Y-\mu)+\sigma^4t^2-\sigma^4t^2\\
&= (Y-\mu-\sigma^2t)^2-\sigma^4t^2
\end{aligned}
$$

y entonces

$$
\begin{aligned}
M(Y;t) &= \exp\left\{t\mu+\frac{\sigma^2t^2}{2}\right\}\int_{-\infty}^{\infty}\underbrace{\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{-[Y-(\mu+\sigma^2t)]^2}{2\sigma^2}\right\}}_{\text{densidad normal}_{(\mu+\sigma^2t,\sigma^2)}}dy\\
&= \exp\left\{t\mu+\frac{\sigma^2t^2}{2}\right\}
\end{aligned}
$$

Si se resta la media a la v.a. Y entonces

$$M(Y-\mu;t)=\exp\left\{-\frac{1}{2}\sigma^2t^2\right\}.$$

Ejemplo 2.2.2 (Familia Exponencial)

Sea Y una v.a. tal que su función de densidad es

$$f_Y(y)=\exp\{y\theta-k(\theta)+a(y)\},$$

entonces su función generadora de momentos es,

$$\begin{aligned}
M(Y;t) &= E(e^{tY}),\\
&= \int_{\mathbb{R}} e^{ty} f_Y(y)dy,\\
&= \int_{\mathbb{R}} e^{ty} e^{(y\theta-k(\theta)+a(y))}dy,\\
&= e^{-k\theta}\int_{\mathbb{R}} e^{y(t+\theta)+a(y)}dy,
\end{aligned}$$

si se completa la intergral con $e^{-k(\theta+t)}$ se obtiene la función de densidad en $(t+\theta)$, y si ésta se integra sobre todo su recorrido el resultado es igual a uno, es decir,

$$\begin{aligned}
M(Y;t+\theta) &= \exp\{-k(\theta)\}\exp\{k(t+\theta)\}\\
&\quad \int_{\mathbb{R}}\exp\{y(t+\theta)-k(\theta+t)+a(y)\}\,dy\\
&= \exp\{-k(\theta)\}\exp\{k(t+\theta)\}\int_{\mathbb{R}} f_Y(y)dy\\
&= \exp\{k(t+\theta)-k(\theta)\}.
\end{aligned}$$

Los momentos no son las únicas constantes que pueden dar una descripción acerca del comportamiento de una distribución, existen también para este propósito otras, las cuales serán descritas a continuación.

2.2.2 Cumulantes y su función generadora.

Es necesario dar una descripción de este nuevo tipo de constantes que también dan una idea acerca de la forma distribucional de una variable aleatoria, pues surgen de manera natural en el desarrollo que se hará en el capítulo 3.

La función generadora de cumulantes de una v.a. Y, se define como:

$$K(Y;t) = \log(M(Y;t)), \tag{2.2.11}$$

y

$$\begin{aligned} K(Y-\mu;t) &= \log(M(Y-\mu;t)) \\ &= \log\left(E(e^{t(Y-\mu)})\right) \\ &= \log\left((e^{-\mu t}E(e^{tY}))\right) \\ &= -t\mu + \log[M(Y;t)] \\ &= -t\mu + K(Y;t). \end{aligned} \tag{2.2.12}$$

El r-ésimo cumulante k_r de la v.a. Y *se define* como el coeficiente de $\frac{t^r}{r!}$ en la expansión en serie de Taylor alrededor de cero de la expresión (2.2.11), entonces

$$\sum_{r=1}^{\infty} k_r \frac{t^r}{r!} = \log(M(Y;t)), \tag{2.2.13}$$

donde k_r es el cumulante de orden r.

Tomando la expansión de (2.2.11) alrededor de cero se tiene

$$\log(M(Y;t)) = \sum_{r=0}^{\infty} \frac{\partial^r}{\partial t^r}(\log M(Y;t))\,|_{t=0}\, \frac{t^r}{r!}$$

Antes de proceder al desarrollo de cada término de la expresión anterior es necesario recordar que

$$\left[\frac{\partial^r M(Y;t)}{\partial t^r}\right]_{t=0} = E(Y^r)$$

por lo tanto,

$$\begin{aligned}
\log[M(Y;t)]_{t=0} &= \left[\log(E(e^{ty}))\right]_{t=0} + \left[\frac{M'(Y;t)}{M(Y;t)}\right]_{t=0}\frac{t}{1!} + \\
&+ \left[\frac{M''(Y:t)}{M(Y;t)} - \frac{(M'(Y;t))^2}{(M(Y;t))^2}\right]_{t=0}\frac{t^2}{2!} + \cdots,
\end{aligned}$$

donde finalmente

$$\log[M(Y;t)_{t=0}] = 1 + E(Y)\frac{t}{1!} + Var(Y)\frac{t^2}{2!} + \cdots,$$

dado que $M(Y;t)\mid_{t=0} = E(e^{tY})\mid_{t=0} = E(e^0) = 1$.

De esta forma, si se sigue desarrollando cada término, se pueden obtener todos los cumulantes en término de los momentos hasta el orden que se desee, lo que implica que

$$\begin{aligned}
k_1 &= \mu_1 \\
k_2 &= \mu_2 = \sigma^2 \\
k_3 &= \mu_3 \\
k_4 &= \mu_4 - 3\mu_2^2 \\
k_5 &= \mu_5 - 10\mu_3\mu_2 \\
k_6 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3 \\
&\vdots
\end{aligned} \tag{2.2.14}$$

Para obtener las relaciones inversas, es decir, los momentos en términos de los cumulantes, se hace directamente de las expresiones anteriores, obteniendo

$$\begin{aligned}
\mu_1 &= k_1 \\
\mu_2 &= k_2 \\
\mu_3 &= k_3 \\
\mu_4 &= k_2 + 3\mu_2^2
\end{aligned}$$

$$
\begin{aligned}
&= k_4 + 3k_2^2 \\
\mu_5 &= k_5 + 10\mu_3^2 - 30\mu_2^2 \\
&= k_5 + k_3k_2 \\
\mu_6 &= k_6 + 15(k_4k_2 + 3k_2^3) + 10k_3^2 - 30k_2^3. \\
&\vdots
\end{aligned}
\tag{2.2.15}
$$

Una vez establecida la relación entre los cumulantes y momentos, se procederá a establecer ciertas propiedades útiles de los cumulantes. Nótese por (2.2.12) que el cumulante k_r, con $r > 0$ es invariante bajo la suma de una constante a la v.a. Y, por tal motivo se pueden considerar como funciones de los momentos alrededor de la media.

Además de esta invarianza ante la suma, debe notarse que para cualquier constante a se tiene

$$k_r(aY) = a^r k_r(Y), \tag{2.2.16}$$

ya que

$$K(aY;t) = \log M(aY;t) = \log E(e^{taY}) = \log E(e^{t'Y})$$

donde $t' = ta$, y por lo visto anteriormente en (2.2.13) se tiene

$$
\begin{aligned}
\log E(e^{t'Y}) &= \sum_{r=0}^{\infty} k_r \frac{t'^r}{r!} \\
&= \sum_{r=0}^{\infty} k_r \frac{(at)^r}{r!} \\
&= \sum_{r=0}^{\infty} a^r k_r \frac{t^r}{r!} \\
&= \sum_{r=0}^{\infty} k_r' \frac{t^r}{r!}
\end{aligned}
$$

con $k_r' = a^r k_r(Y)$, lo que implica que

$$k_r' = k_r(aY).$$

Más aun, k_r es una función de los momentos solamente en un orden no superior a r. por esto, puede ser definido sin considerar la convergencia de los momentos para órdenes mayores a r.

Ejemplo 2.2.3 (Distribución Normal).

Sea Y una v.a. con distribución normal con media μ y varianza σ^2, del *Ejemplo 2.2.1* se tiene que

$$M(Y;t) = \exp\left\{t\mu + \frac{1}{2}t^2\sigma^2\right\},$$

$$M(Y-\mu;t) = \exp\left\{\frac{1}{2}t^2\sigma^2\right\},$$

y su correspondiente función generadora de cumulantes es

$$K(Y;t) = \log M(Y;t) = t\mu + \frac{1}{2}t^2\sigma^2,$$

$$K(Y-\mu;t) = \log M(Y-\mu;t) = \frac{1}{2}\sigma^2 t^2,$$

De estas dos funciones generadoras de cumulantes se observa que para el caso de una distribución normal con parámetros μ, σ^2, el cumulante de orden uno, $k_1 = \mu$, el cumulante de orden dos, $k_2 = \sigma^2$ y $k_r = 0$ para toda $r \geq 3$. De aquí que si en algún estudio se obtiene que una v.a. tiene $k_r = 0$ para toda $r \geq 3$, podría pensarse que tiene una distribución normal o casi es una v.a. normal. Lamentablemente no se podría decir lo mismo si se obtienen los momentos, ya que para el caso más simple de una distribución $N(0,\sigma^2)$, como se vió en el *Ejemplo 2.2.1*, el n-ésimo momento se calcula como

$$E(X^n) = \begin{cases} 0 & \text{si } n \text{ es impar} \\ \frac{\sigma^n n!}{2^{n/2}\left(\frac{n}{2}!\right)} & \text{si } n \text{ es par} \end{cases}$$

Ejemplo (2.2.4) (Familia Exponencial).

Si Y tiene una densidad dada por $\exp\{y\theta - M(\theta) + a(y)\}$, entonces su función generadora de momentos es, por el *ejemplo (2.2.2)*

$$M(Y;t) = \exp\{M(\theta+t) - M(\theta)\},$$

y su correspondiente función de cumulantes es

$$K(Y;t) = \log M(Y;t) = M(\theta+t) - M(\theta),$$

Ejemplo (2.2.5) (Densidad exponencial).

$$K(Y;t) = \log M(Y;t) = \log\left(\frac{\lambda}{\lambda - t}\right) = \log(\lambda) - \log(\lambda - t), \qquad (2.2.17)$$

Otra propiedad de los cumulantes, que proporciona una manera más fácil de trabajar, es la de poder ser escritos sin dimensiones, es decir, el r-ésimo cumulante estandarizado se define como

$$\rho_k = \frac{k_r}{\sigma^r}, \qquad (2.2.18)$$

para $r = 3, 4, \cdots$,

en particular $\rho_3 = \frac{k_3}{\sigma^3}$ y $\rho_4 = \frac{k_4}{\sigma^4}$, son denotados frecuentemente por γ_1 y γ_2, que representan una medida de asimetría (skewness) y de agudeza(kurtosis) respectivamente de la función de densidad.

Hasta el momento se ha analizado que los momentos, de cierto orden están relacionados a cumulantes del mismo orden, siendo estos últimos los de mayor interés, ya que se utilizarán en el capítulo posterior.

Dado que los cuatro primeros momentos o cumulantes son de gran utilidad para proporcionar el grado de variabilidad de la información con la que se cuenta para ciertos estudios, se procederá a dar una breve descripción de cada uno de ellos.

(i) Esperanza.

El valor esperado, es una medida de tendencia central, la cual dice donde está situado el centro de masa de la distribución de probabilidades de una variable aleatoria. Es también el valor promedio de una variable aleatoria si el mismo experimento aleatorio se repite una y otra vez.

Para ejemplificar el concepto supóngase que se cuenta con la siguiente información.

La probabilidad de que una casa de cierto tipo quede destruida por un incendio en cualquier período de doce meses es 0.005. Una compañía de seguros ofrece vender al propietario de tal casa una póliza contra incendio por 20,000.00 pesos, el tiempo de cobertura es igual a un año por una prima de 150.00 pesos. La ganancia G para la compañía es una variable aleatoria con valores posibles de 150.00 pesos si la casa no se incendia, y de -19,850.00 pesos si se incendia durante el año cubierto por la póliza. La función de densidad de G es $f(g_i)$, donde g_i es la ganancia o pérdida asociada, por lo tanto,

$$E(G) = 150(0.995) + (-19,850)(0.005) = 50.$$

Desde el punto de vista de la compañía de seguros, la ganancia esperada por una póliza de seguro debe ser positiva para permitir a la aseguradora pagar los costos de administración y acumular reservas para pagar a sus beneficiarios y tenedores de pólizas. Pero desde el punto de vista del comprador, el seguro, como cualquier juego de azar que se hace para obtener una utilidad, tiene un valor esperado negativo.

(ii) Desviación estándar.

La varianza y la desviación estándar son frecuentemente empleadas como medidas de variabilidad, teniendo ambas interpretaciones intuitivas.

Cuando se comparan uno o más conjuntos de datos cuyas unidades de medición son idénticas, se puede decir que una muestra tiene un grado de dispersión menor que otra si la primera tiene una varianza menor, sin embargo, si esta aceveración se hace acerca de un sólo conjunto de datos, de los cuales se obtenga el mismo grado de dispersión no podrá ser especificado el grado de variabilidad. Un concepto que proporciona más claridad acerca de las comparaciones entre conjuntos de datos es el *Coeficiente de Variación*, el cual consiste en dividir la desviación estándar entre la media, esto es, $cv = \frac{\sigma}{\mu}$. Para ejemplificarlo, supóngase que se cuenta con la siguiente información:

El peso promedio de una muestra de elefantes es 24,000 libras y su desviación estándar es de 1,285 libras, en tanto que para un conjunto de ratones el peso promedio es igual a 1.05 libras con una desviación estándar de 0.16 libras.

Claramente se puede ver que el peso de los elefantes es mucho mayor que el de los ratones, sin embargo, el peso de los elefantes es más pequeño como un porcentaje de su propia media que el de los ratones, es decir, el coeficiente de variación para los elefantes es de 5.4 por ciento, en tanto que el de los ratones es de 15.2 por ciento.

Con esto se puede observar que la variabilidad del peso de los ratones es aproximadamente tres veces mayor que la variabilidad del peso de los elefantes apesar de que la desviación estándar de estos últimos es mayor.

(iii) Medida de simetría.(Skewness)

Este concepto cobra gran importancia debido a que las consideraciones teóricas en la inferencia estadística frecuentemente están basadas en la suposición de que las poblaciones se distribuyen como *normales*, sirviendo esta medida para conocer o prevenir los errores en los que posiblemente se caiga cuando se hacen este tipo de suposiciones.

La medida de asimetría es conocida como medida de *Pearson*, la cual está basada en la relación que existe entre la media, la moda y la mediana. Debe recordarse que para una distribución simétrica estas tres medidas son iguales, pero en una distribución asimétrica la media se mueve alrededor de la moda en la dirección que indique la simetría, y la mediana se encontrará, como su nombre lo indica, en medio.

Consecuentemente, la distancia entre la media y la moda puede ser usada para medir el grado de simetría de una distribución, la cual será más positiva o más negativa dependiendo de que tan asimétrica sea la distribución. Precisando, se tiene que: *asimetría = media - moda*. Cuanto mayor es esta

distancia, negativa o positiva, tanto más asimétrica es la distribución.

Es necesario mencionar que esta medida tiene dos defectos en su aplicación, primero es una medida absoluta, es decir, el resultado está expresado en términos de las unidades originales de la distribución y esto provoca cambios cuando las unidades de medición cambian, y en segundo, la misma medición de asimetría tiene un diferente significado para diferentes series con diferentes grados de variabilidad. Para eliminar estos defectos, puede introducirse una medida relativa de asimetría. Esto se logra por *el coeficiente pearsoniano de asimetría*, designado por sk_p, y expresado simbólicamente como

$$sk_p = \frac{\bar{X} - X_m}{\hat{s}} = \gamma_1, \tag{2.2.19}$$

donde $\bar{X}$ = media, X_m = moda y $\hat{s}$ = desviación estándar.

La aplicación de (2.2.19) supone otra dificultad que surge porque el valor modal de muchas distribuciones sólo es una aproximación, pero el lugar de la mediana se encuentra más fácilmente. Se sabe que en distribuciones moderadamente asimétricas[1] se verifica la relación

$$X_m \doteq \bar{X} - \left[\bar{X} - 3(\bar{X} - X_{.5})\right] = 3(\bar{X} - X_{.5}).$$

Con este resultado, la ecuación (2.2.19) puede escribirse de la siguiente forma

$$sk_p = \frac{3(\bar{X} - X_{.5})}{\hat{s}}. \tag{2.2.20}$$

Por esta razón $sk_p = 0$ para una distribución asimétrica, negativa para una distribución simétrica a la izquierda y positiva para una distribución asimétrica a la derecha. Más precisamente esta medida variará dentro de los límites de ± 3; sin embargo, sólo en raras ocasiones el valor de sk_p supera los límites de ± 1.

Es necesario recordar que el coeficiente

$$\frac{\mu_3}{\sigma^3} = \rho_3 = \gamma_1,$$

[1] Ref. [7] Sección 3.5

se conoce como el coeficiente de asimetría y puede ser estimado descriptivamente en términos de la muestra como se mostró anteriormente en (2.2.20).

(*iv*) **Grado de agudeza (Kurtosis).**

Se define como una medida de exceso, la cual indica el grado de agudeza o grosor de una densidad con respecto a su centro.

Los valores positivos de $\frac{\mu_4}{\sigma^4} - 3$, que es conocido como *coeficiente de kurtosis*, son usados en algunas ocasiones para indicar que una densidad es más aguda alrededor de su centro que la densidad de una curva normal, en tanto que los valores negativos son usados para indicar que una densidad es más ancha alrededor de su centro que la densidad de una curva normal.

1. **Si $\frac{\mu_4}{\sigma^4} > 3$, entonces la distribución es más aguda que una distribución normal.**
2. **Si $\frac{\mu_4}{\sigma^4} < 3$, entonces la distribución es más ancha que una distribución normal.**
3. **Si $\frac{\mu_4}{\sigma^4} = 3$, entonces se dice que la distribución tiene la agudeza de una distribución normal.**

Para estimar descriptivamente el coeficiente de kurtosis, basta calcular

$$k = \frac{\frac{1}{2}(x_{.75} - x_{.25})}{x_{.90} - x_{.10}},$$

que representa la proporción entre el punto medio de la distancia que existe entre el cuantil $x_{.75}$ y $x_{.25}$, y la distancia entre los cuantiles $x_{.90}$ y $x_{.10}$.

Obsérvese que para el caso (1), la diferencia entre las dos distancias tiende a ser muy pequeña, lo que indica que la información está concentrada alrededor del cuantil $x_{.50}$, es decir, el coeficiente k, se aproxima a 0.5 dependiendo de la agudeza de la curva. En contraste, cuanto más plana es la distribución, caso (2), tanto más amplia es la longitud entre $x_{.90}$ y $x_{.10}$, por lo que si esta

diferencia tiende a infinito, entonces $k = 0$, es decir se tiene una curva completamente plana. Para el caso (3), es evidente que el coeficiente k toma valores cercanos a 0.25, ya que la curva a considerar es una *Normal.*

Para el cálculo de los cuantiles $x_{.5}$, $x_{.25}$, $x_{.75}$, $x_{.90}$ y $x_{.10}$, que son utilizados para determinar los coeficientes de asimietría y de kurtosis, es necesario introducir el ***método de interpolación algebraica***, y con ello poder hallar sus estimadores a través de medidas descriptivas para datos agrupados.

La fórmula general de interpolación algebraica para el p-ésimo percentil entre datos agrupados es:

$$X_{p} = L_{p} + \left[\frac{\frac{p(n+1)}{100} - F_{p}}{f_{p}} \right] c. \qquad (2.2.21)$$

donde:

p = porcentaje

L_{p} = límite verdadero de clase inferior de la clase en donde se encuentra el p-ésimo percentil.

F_{p} = frecuencia acumulada en la clase que precede inmediatamente a la clase en donde se encuentra el p-ésimo percentil.

f_{p} = frecuencia de clase en la clase que precede inmediatamente a la clase en donde se encuentra el p-ésimo percentil.

c = longitud del intervalo de clase de la clase en donde se encuentra el p-ésimo porcentil.

Además de estimar descriptivamente, el coeficiente de kurtosis y de asimetría, tal y como se mencionó anteriormente, existe también la posibilidad de estimar dichos coeficientes en términos de las observaciones muestrales de v.a.i.i.d, es decir

$$\frac{\bar{\mu}_3}{\sigma^3} = \frac{E\left[(X-\mu)^3\right]}{\sigma^3} = \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{n}}{s^3}. \qquad (2.2.22)$$

y

$$\frac{\hat{\mu}_4}{\sigma^4} - 3 = \frac{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n}}{s^4} - 3, \tag{2.2.23}$$

denotados por los estimadores muestrales de los coeficientes de simetría y kurtosis respectivamente.

Con esta breve explicación se ha terminado de analizar a los momentos, a los cumulantes y la relación entre ellos. Se procede entonces a retomar este concepto, pero para el caso de sumas de variables aleatorias.

2.3 Momentos, cumulantes y funciones generadoras para sumas de variables aleatorias.

En la sección 1 de este capítulo se discutió sobre la obtención de la distribución exacta de la suma parcial S_n de $Y_1, Y_2, \ldots, Y_n$ v.a.i.i.d. copias de una v.a. Y. Se observó que a medida que n crece el cálculo de la distribución se complica cada vez más, llevando a la necesidad de una aproximación de tipo numérico. Por tal motivo se ha decidido hacer el análisis de la distribución de S_n en el límite, con lo cual se necesitarán definir las funciones generadoras de momentos y cumulantes de S_n.

Recordando que $E(S_n) = n\mu$ y $Var(S_n) = n\sigma^2$, la función generadora de momentos de S_n se calcula como,

$$\begin{aligned} M(S_n;t) &= E(e^{tS_n}) \\ &= E\left(\exp\left[t\sum_{j=1}^{n} Y_j\right]\right) \\ &= \prod_{j=1}^{n} E(e^{tY_j}) \end{aligned}$$

ya que las Y_i's son independientes y con la misma distribución, por lo tanto.

$$M(S_n;t) = (M(Y;t))^n, \tag{2.3.24}$$

y su función generadora de cumulantes como

$$\begin{aligned} K(S_n;t) &= \log M(S_n;t) \\ &= \log M(Y;t)^n \\ &= n\log M(Y;t) \\ &= nK(Y;t). \end{aligned} \tag{2.3.25}$$

De esta expresión se tiene que el r-ésimo cumulante de S_n está dado por

$$k_r(S_n) = nk_r(Y), \tag{2.3.26}$$

para $r = 1, 2, 3, \cdots$, expresión que generaliza la ecuación (2.1.3).

Debe notarse la simplicidad de la expresión (2.3.26), razón suficiente para decidir trabajar con el concepto de cumulantes, además de que $\mu_r(S_n) \neq n\mu_r(Y)$ para $r > 3$.

Por otro lado debe notarse que debido a la invarianza bajo transformaciones lineales se tiene que

$$\rho_r(S_n^*) = \rho_r(S_n). \tag{2.3.27}$$

Lo anterior se observa del hecho de que

$$\begin{aligned} M(S_n^*;t) &= E\left(\exp\left[\sum_{i=1}^n t\left(\frac{Y_i - n\mu}{\sigma\sqrt{n}}\right)\right]\right) \\ &= \prod_{i=1}^n E\left(\exp\left[\frac{tY_i}{\sigma\sqrt{n}}\right]\exp\left[-\frac{n\mu t}{\sigma\sqrt{n}}\right]\right) \\ &= \exp\left[-\frac{n\mu t}{\sigma\sqrt{n}}\right]\prod_{i=1}^n E\left(\exp\left[\frac{tY_i}{\sigma\sqrt{n}}\right]\right) \\ &= \exp\left[-\frac{n\mu t}{\sigma\sqrt{n}}\right] M\left(Y;\frac{t}{\sigma\sqrt{n}}\right)^n, \end{aligned} \tag{2.3.28}$$

y que

$$\begin{aligned} K(S_n^*;t) &= \log M(S_n^*;t) \\ &= \log\left[\left(\exp\left[\frac{-n\mu t}{\sigma\sqrt{n}}\right]\right)\left(M\left(Y;\frac{t}{\sigma\sqrt{n}}\right)\right)^n\right] \\ &= \frac{-n\mu t}{\sigma\sqrt{n}} + nK(Y;\frac{t}{\sigma\sqrt{n}}). \end{aligned} \tag{2.3.29}$$

De aquí que por (2.2.11),

$$\begin{aligned} K(S_n^*,t) &= \frac{-n\mu t}{\sigma\sqrt{n}} + nK\left(Y;\frac{t}{\sigma\sqrt{n}}\right) \\ &= \underbrace{K\left(Y-\frac{\mu}{\sigma\sqrt{n}};\frac{t}{\sigma\sqrt{n}}\right)+\cdots+K\left(Y-\frac{\mu}{\sigma\sqrt{n}};\frac{t}{\sigma\sqrt{n}}\right)}_{\text{n veces}} \\ &= nK\left(Y-\frac{\mu}{\sigma\sqrt{n}};t'\right), \end{aligned}$$

con $t' = \frac{t}{\sigma\sqrt{n}}$.

Luego entonces

$$\begin{aligned} K(S_n,t) &= nK(Y;t) \\ K(S_n^*,t) &= nK\left(Y-\frac{\mu}{\sigma\sqrt{n}};t'\right) \end{aligned}$$

y de aquí que, por la propiedad de invarianza ante la suma de una constante,

$$k_r(S_n) = k_r(S_n^*)$$

por lo tanto

$$\frac{k_r(S_n)}{\sigma^r} = \frac{k_r(S_n^*)}{\sigma^r}.$$

Además debe notarse que,

$$\begin{aligned}\rho_r(S_n) &= \frac{k_r(S_n)}{(\sigma_{s_n})^r} \\ &= \frac{k_r(S_n)}{(\sqrt{n}\sigma)^r}\end{aligned}$$

por (2.3.26) se tiene

$$\rho_r(S_n) = \frac{nk_r(Y)}{(\sqrt{n}\sigma)^r},$$

así

$$\begin{aligned}\rho_3(S_n) &= \frac{nk_3(Y)}{n^{3/2}\sigma^3} = \frac{\rho_3(Y)}{\sqrt{n}} \\ \rho_4(S_n) &= \frac{nk_4(Y)}{n^2\sigma^4} = \frac{\rho_4(Y)}{n^2} \\ &\vdots \\ \rho_r(S_n) &= \frac{nk_r(Y)}{n^{r/2}\sigma^r} = \frac{\rho_r(Y)}{n^{(r/2)-1}},\end{aligned} \tag{2.3.30}$$

y obviamente también para $\rho_r(S_n^*)$ se tiene la igualdad descrita en (2.3.31), esto es,

$$\rho_r(S_n^*) = \frac{\rho_r(Y)}{n^{(r/2)-1}}, \tag{2.3.31}$$

lo cual será de utilidad para la expansión en series de Edgeworth que se desarrollará en el capítulo 3.

2.4 Orden de Convergencia.

Antes de continuar, es importante conocer o recordar el concepto de orden convergencia, ya que con ello se podrán analizar expresiones que surgirán más adelante en este trabajo.

Definición 2.4.1 *Si U y V son funciones reales de una variable aleatoria real, donde $V \neq 0$ para toda x, se tiene que:*

1. $u(x) = \mathcal{O}(v(x))$ *cuando* $x \to x_0$, *si* $\left|\frac{u(x)}{v(x)}\right| < M$ *con* $x \to x_0$,

2. $u(x) = o(v(x))$ *cuando* $x \to x_0$, *si* $\lim_{x \to x_0} \frac{u(x)}{v(x)} = 0$.

3. $u(x) \sim v(x)$ *cuando* $x \to x_0$, *si* $\lim_{x \to x_0} \frac{u(x)}{v(x)} = 1$.

Ejemplo 2.4.1

Si $f(n) = \frac{a}{n} + \frac{b}{n^2} + \frac{c}{n^3}$ entonces $\lim_{n \to \infty} f(n) = 0$, es decir

$$\lim_{n \to \infty} \left(\frac{\frac{a}{n} + \frac{b}{n^2} + \frac{c}{n^3}}{1} \right) = 0$$

entonces

$$f(n) = o(1).$$

Además

$$f(n) = \frac{a}{n} + \frac{b}{n^2} + \mathcal{O}(n^{-3}),$$

ya que

$$\left| \frac{\frac{c}{n^3}}{\frac{1}{n^3}} \right| = c < M,$$

esto implica que $\frac{c}{n^3}$ se acota en el orden n^{-3}.

Análogamente se tiene que

$$f(n) = \frac{a}{n} + \mathcal{O}(n^{-2}),$$

ya que

$$\left| \frac{\frac{b}{n^2}}{\frac{1}{n^2}} \right| = b,$$

y también

$$\lim_{n \to \infty} \frac{\frac{b}{n^2}}{\frac{1}{n}} = \lim_{n \to \infty} \frac{b}{n} = 0.$$

entonces

$$f(n) = \frac{a}{n} + o(n^{-1}).$$

Finalmente,

$$f(n) \sim \frac{a}{n}$$

ya que

$$\lim_{n\to\infty} \frac{\frac{a}{n}}{\frac{a}{n}} = 1.$$

De esta forma se puede observar que si en el límite una función se acota al dividir entre otra función de un orden específico, entonces también se podrá conseguir que dicha función tienda a cero dividiendo entre otra de un orden mayor al anterior, retomando el *ejemplo (2.4.1)* $\frac{1}{n^2} = O(n^{-2})$ ya que $|\frac{n^2}{n^2}| < M$, pero también $\frac{1}{n^2} = o(n^{-1})$ ya que $\lim_{n\to\infty} \frac{1}{n} = 0$.

De manera general, supóngase $f(Y_1, Y_2, \ldots, Y_n) = k + o(1) \geq k$ con $k \in \mathbb{R}$ y n lo suficientemente grande. Suponga que en este caso $o(1) = \frac{c}{n}$, donde c es una constante que está en los reales. Además es importante notar que la aproximación de $f(Y_1, Y_2, \ldots, Y_n)$ a k sería más exacta si se tuvieran órdenes del tipo

$$\begin{aligned} O(n^{-\frac{1}{2}}) &= o(n^{-\frac{1}{2}}) = \frac{c}{n^{\frac{1}{2}}} \\ O(n^{-\frac{2}{2}}) &= o(n^{-\frac{2}{2}}) = \frac{c}{n^{\frac{2}{2}}} \\ O(n^{-\frac{3}{2}}) &= o(n^{-\frac{3}{2}}) = \frac{c}{n^{\frac{3}{2}}}, \end{aligned}$$

ya que cuanto mayor es el exponente de n mayor es el denominador y la expresión tenderá más rápido a cero sin la necesidad de que n sea tan grande,(n es el tamaño de la muestra). Con este concepto, se puede llegar a establecer el orden en el que converge la expresión (2.3.31), es decir,

$$\left| \frac{\frac{k_r(Y)}{\sigma^r n^{\frac{r}{2}-1}}}{\frac{1}{n^{\frac{r}{2}-1}}} \right| = \frac{k_r(Y)}{\sigma^r} < M,$$

lo que implica que

$$\rho_r(S_n) = \mathcal{O}(n^{-\frac{r}{2}+1}), \tag{2.4.32}$$

además si

$$\frac{\frac{k_r(Y)}{\sigma^r n^{\frac{r}{2}-1}}}{\frac{1}{n^{\frac{(r-1)}{2}-1}}} = \frac{k_r(Y) n^{\frac{(r-1)}{2}-1}}{\sigma^r n^{\frac{r}{2}-1}},$$

lo que implica que

$$\lim_{n\to\infty} \frac{k_r(Y)}{\sigma^r n} = 0,$$

por lo tanto

$$\rho_r(S_n) = o(n^{\frac{(r-1)}{2}-1}). \tag{2.4.33}$$

Además de de las propiedades de convergencia vistas anteriormente, se tiene

$$\begin{aligned}
\mathcal{O}(n^a)\mathcal{O}(n^b) &= \mathcal{O}(n^{a+b}) \\
\mathcal{O}(n^a)o(n^b) &= o(n^{a+b}) \\
o(n^a)o(n^b) &= o(n^{a+b}).
\end{aligned}$$

Por ejemplo,

$$\begin{aligned}
n^a &= \mathcal{O}(n^a)\text{ , } n^b = \mathcal{O}(n^b) \text{ y esto implica que } n^{a+b} = \mathcal{O}(n^{a+b}) \\
n^a &= \mathcal{O}(n^a)\text{ , } n^{b-1} = o(n^b) \text{ y esto implica que } n^{a+b-1} = o(n^{a+b})
\end{aligned}$$

Con esto, puede decirse a *grosso modo*, que o domina a $\mathcal{O}$.

2.5 Momentos, cumulantes y funciónes generadora para sumas aleatorias de variables aleatorias.

Sumas de la forma $Y_1 + \ldots + Y_N$, donde N es una v.a., pueden encontrarse de manera natural en muchos ejemplos, como los que se mostrarán en las siguientes secciones.

a) Teoría de colas.

Sea N el número de clientes que llegan a alguna tienda de servicios en un período de tiempo específico, y sea Y_i el tiempo de servicio requerido para el i-ésimo cliente. entonces $S_N = Y_1 + \ldots + Y_N$, es la demanda total por tiempo de servicio.

b)Teoría de Riesgo.

Supóngase que N es el total de reclamaciones que recibe una compañía de seguros en una semana. Sea Y_i el monto de la i-ésima reclamación, entonces la cantidad total a la que la compañía debe hacer frente es, $S_N = Y_1 + \ldots + Y_N$.

c) Modelos poblacionales.

Sea N el número de plantas de cierta especie en una localidad específica, y sea Y_i el número de semillas producidas por la i-ésima planta, entonces $S_N = Y_1 + \ldots + Y_N$, es el número total de semillas producidas en la localidad.

La definición de sumas aleatorias de variables aleatorias que aquí se proporcionará es muy débil debido a los supuestos que se establecen, sin embargo es la única que será necesaria definir en este trabajo.

Definición 2.5.1 *Sean $\{Y_1, Y_2, \ldots\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas y sea N una v.a. discreta independiente de $Y_1, Y_2, \ldots$ con función de probabilidad $P_N(n) = P\{N = n\}$ para $n = 0, 1, \ldots$, se define la suma aleatoria parcial como*

$$S_N = \begin{cases} 0 & si\ N = 0 \\ Y_1 + Y_2 + \cdots + Y_N & si\ N > 0 \end{cases} \tag{2.5.34}$$

Una vez establecida la definición se procederá a hacer el cálculo de momentos de sumas aleatorias de variables aleatorias.

Supóngase que Y_k y N tienen momentos finitos, entonces

$$E(Y_k) = \mu$$

$$\begin{aligned} Var(Y_k) &= \sigma^2 \\ E(N) &= \nu \\ Var(N) &= \tau^2. \end{aligned}$$

Si se determina la media y la varianza para $S_N = Y_1 + Y_2 + \cdots + Y_N$ se tiene que

$$E(S_N) = \mu\nu \tag{2.5.35}$$

$$Var(S_N) = \nu\sigma^2 + \mu^2\tau^2. \tag{2.5.36}$$

Para la demostración de (2.5.35) y (2.5.36) es necesario enunciar el concepto esperanza condicional y algunas de sus propiedades.

Esperanza Condicional.

Sea g una función para la cual la esperanza de $g(X)$ es finita. Se define el valor de la esperanza condicional de $g(X)$ dado $Y = y$ como,

$$E(g(X)|Y = y) = \sum_y g(X)P_{X|Y}(x|y) \quad \text{si} \quad P_Y(y) > 0,$$

ya que la media condicional no está definida en valores de Y para las cuales $P_Y(y) = 0$.

La ley de probabilidad total para la esperanza condicional se lee como

$$E(g(X)) = \sum_y E(g(X)|Y = y)P_Y(y). \tag{2.5.37}$$

El valor de la esperanza condicional $E[g(X)|Y = y]$ es una función de la variable real y. Si esta función es evaluada en una v.a. Y, es decir se toma la composición con Y, se obtiene la v.a. que se denota por $E[g(X)|Y]$, entonces (2.5.37) se puede reescribir como

$$E(g(X)) = E(E(g(X)|Y)) = \sum_y E(g(X)|Y = y)P_Y(y).$$

A continuación, se proporcionarán algunas propiedades de esperanza condicional.

Sean X y Y variables aleatorias distribuídas conjuntamente, c un número real. g una función para la cual $E(|g(X)|) < \infty$. h una función acotada; y v una función de dos variables para la cual $E(|v(X,Y)|) < \infty$, entonces las propiedades de esperanza condicional son:

1. $E(c_1 g_1(X_1) + c_2 g_2(X_2)|Y) = c_1 E(g_1(X_1)|Y) + c_2 E(g_2(X_2)|Y)$.
2. Si $g \geq 0$ entonces $E(g(X)|Y = y) \geq 0$.
3. $E(v(X,Y)|Y = y) = E(v(X,y)|Y = y)$.
4. $E(g(X)|Y = y) = E(g(X))$, si X y Y son independientes
5. $E(g(X)h(Y)|Y = y) = h(y)E(g(X)|Y = y)$ y

$$\begin{aligned} E(g(X)h(Y)) &= \sum_Y h(y)E(g(X)|Y = y)P_Y(y), \\ &= E(h(Y)E(g(X)|Y)). \end{aligned}$$

como consecuencia de (1) y (5), para $g \equiv 1$ y $h \equiv 1$ se obtiene,

6. $E(c|Y = y) = c$.
7. $E(h(Y)|Y = y) = h(y)$.
8. $E(g(X)) = \sum_y E(g(X)|Y = y)\, P_Y(y) = E(E(g(X)|Y))$.

Ahora se procederá a la demostración de (2.5.35), que está basada en las propiedades de esperanza condicional.

Demostración.

Por (2.5.37) se tiene que,

$$E(S_N) = \sum_{n=1}^{\infty} E(Y_1 + Y_2 + \cdots + Y_N \mid N = n)P_N(n)$$

por el punto (4)

$$
\begin{aligned}
E(S_N) &= \sum_{n=1}^{\infty} E(Y_1+Y_2+\cdots+Y_N|N=n)P_n(n)\\
&= \sum_{n=1}^{\infty} E(Y_1+Y_2+\cdots+Y_n)P_N(n)\\
&= \sum_{n=1}^{\infty} \mu n\, P_N(n)\\
&= \mu\sum_{n=1}^{\infty} n\, P_N(n)\\
&= \mu E(N)\\
&= \mu\nu.
\end{aligned}
$$

□

Para la demostración de (2.5.26) se tiene.

$$
\begin{aligned}
Var(S_N) &= E(S_N-\mu\nu)^2\\
&= E((S_N-N\mu+N\mu-\mu\nu)^2)\\
&= E(S_N-N\mu)^2+E(\mu^2(N-\nu)^2)+2E(\mu(S_N-N\mu)(N-\nu)).
\end{aligned}
$$

Desarrollando $E((S_N-N\mu)^2)$ se obtiene

$$
\begin{aligned}
E((S_N-N\mu)^2) &= \sum_{n=1}^{\infty} E(Y_1+\cdots+Y_N-N\mu|N=n)^2P_N(n)\\
&= \sum_{n=1}^{\infty} E(Y_1+\cdots+Y_n-n\mu)^2P_N(n)\\
&= \sum_{n=1}^{\infty} Var(Y_1+\cdots+Y_n)P_N(n)\\
&= \sum_{n=1}^{\infty} n\sigma^2P_N(n)\\
&= \sigma^2\sum_{n=1}^{\infty} nP_N(n)\\
&= \sigma^2E(N)\\
&= \sigma^2\nu
\end{aligned}
$$

Desarrollando $E(\mu^2(N-\nu)^2)$ se obtiene

$$\begin{aligned} E(\mu^2(N-\nu)^2) &= \mu^2 E((N-\nu)^2) = \mu^2 Var(N), \\ &= \mu^2\tau^2. \end{aligned}$$

Mientras que $E(\mu(S_N - N\mu)(N-\nu)) = 0$ ya que

$$\begin{aligned} E(S_N - N\mu) &= E(Y_1 + Y_2 + \cdots + Y_n | N = n)P_N(n) \\ &= E(Y_1 + \cdots + Y_n - n\mu)P_N(n) \\ &= (n\mu - n\mu)P_N(n) \\ &= 0 \end{aligned}$$

Por lo tanto

$$Var(S_N) = \sigma^2\nu + \mu^2 + \tau^2.$$

$\square$

Sea $F(y)$ la función de distribución de las v.a. Y_i's, las cuales se suponen independientes, y sea Y una v.a. con esta distribución acumulada, tal que

$$E(Y^k) = p_k ,$$

es el k-ésimo momento alrededor del origen y

$$M(Y;t) = E(e^{tY}),$$

la función generadora de momentos de la v.a. Y. También suponga que se tiene la función generadora de momentos de la v.a. N, la cual está dada por

$$M(N;t) = E(e^{tN}),$$

entonces la función generadora de momentos de la suma aleatoria de variables aleatorias S_n es:

$$M(S_N;t) = E(e^{tS_N}) = E(E(e^{tS_N} \mid N)),$$

si se calcula

$$\begin{aligned} E(e^{tS_N} \mid N=n) &= E(e^{tS_n} |_{N=n}) \\ &= E(\exp\{t(Y_1+Y_2+\ldots+Y_n)\}) \\ &= \prod_{i=1}^{n} E(e^{tY_i}) = (E(e^{tY}))^n, \end{aligned}$$

de aquí que

$$\begin{aligned} E(e^{tS_n} \mid N) &= (E(e^{tY}))^n \circ N = (E(e^{tY}))^N \\ &= [M(Y;t)]^N, \end{aligned}$$

lo que implica que

$$\begin{aligned} M(S_N;t) &= E[M(Y;t)^N] = E[e^{N\log(M(Y;t))}] \qquad (2.5.38) \\ &= M(N; logM(Y;t)). \end{aligned}$$

□

Ejemplo 2.5.1

Suponga que N tiene una distribución geométrica, esto es, la función de densidad de N está dada por

$$P(N=n) = pq^n \text{ , } n = 0,1,2,\ldots$$

donde $0 < q < 1$ y $p = 1-q$. Calcular la $M(S_N;t)$

Solución.

$$M(N;t) = E\left(e^{tN}\right) = \sum_{n=0}^{\infty} p\left(qe^t\right)^n = \frac{p}{1-qe^t} \text{ ,}$$

la fórmula (2.5.39) lleva a que

$$M(S_N;t) = \frac{p}{1-q(\exp\{\log(M(Y;t))\})} = \frac{p}{1-q(M(Y;t))}.$$

Distribución de sumas aleatorias.

Suponga que $Y_1, \ldots, Y_n$, son v.a. continuas con función de densidad de probabilidades $f_Z(z)$. Para $n \geq 1$ la función de densidad de probabilidades de la función compuesta por $Y_1 + Y_2 + \cdots + Y_n$, es la n-ésima convolución de $f_Z(z)$, denotada por $f^{(n)}(z)$, y definida recursivamente por. $f^1(z) = f(z)$ y

$$f^{(n)}(z) = \int f^{(n-1)}(z-\mu)f(u)du.$$

Debido a que $Y_1, \ldots, Y_n$ y N son independientes, se tiene que la función de densidad condicional de $Y_1 + Y_2 + \cdots + Y_n$ dado que $N = n > 1$ es también $f^{(n)}(z)$. Suponga ahora que $P\{N = 0\} = 0$, entonces, por la ley de probabilidad total S_N es continua y su función de densidad es

$$\begin{aligned} f(S_N) &= \sum_{n=1}^{\infty} f_{S_N|N=n}(s_N|N=n)P_N(n) \\ &= \sum_{n=1}^{\infty} f_{S_n}(s_n)P_N(n) \\ f(S_N) &= \sum_{n=1}^{\infty} f^{(n)}(s_n)P_N(n). \end{aligned} \qquad (2.5.39)$$

Nota.

Cuando $N = 0$ con probabilidad positiva, entonces $S_N = Y_1 + \ldots + Y_N$, es una variable que tiene dos componentes en su distribución: continua y discreta. Recordando el supuesto de que $Y_1 + Y_2 + \ldots + Y_k$ son continuas con f.d.p. $f_Z(z)$, entonces

$$P[S_N = 0] = P[N = 0] = P_N(0),$$

mientras que para $a < S_N < b$ con $0 < a < b$, o bien, $a < b < 0$ es

$$P\{a < S_N < b\} = \int_a^b \left[\sum_{n=1}^{\infty} f^{(n)}(z)P_N(n)\right] dz. \qquad (2.5.40)$$

Ejemplo (2.5.1)

Suponga *una suma geométrica de variables aleatorias exponenciales*, entonces

$$f(z) = \begin{cases} \lambda e^{-\lambda z} & \text{para } z \geq 0 \\ 0 & \text{para } z < 0 \end{cases}$$

es una densidad exponencial y

$$P_N(n) = \beta(1-\beta)^{n-1} \text{para } n = 1, 2, \ldots,$$

es una densidad geométrica.

Recordando que la convolución de densidades exponenciales con parámetro λ es una gamma con parámetro (n, λ), se tiene que la densidad para $S_N = Y_1 + \cdots + Y_N$ está dada por

$$\begin{aligned}
f_{S_N}(z) &= \sum_{n=1}^{\infty} f^{(n)}(z) P_N(n) \\
&= \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z} \beta(1-\beta)^{n-1} \\
&= \lambda\beta e^{-\lambda z} \sum_{n=1}^{\infty} \frac{\lambda^{n-1} z^{n-1} (1-\beta)^{n-1}}{(n-1)!} \\
&= \lambda\beta e^{-\lambda z} \sum_{n=1}^{\infty} \frac{(\lambda z(1-\beta))^{n-1}}{(n-1)!} \\
&= \lambda\beta e^{-\lambda z} e^{\lambda(1-\beta)z} \\
&= \lambda\beta e^{-\lambda z + \lambda(1-\beta)z} \\
&= \lambda\beta e^{-\lambda z + (\lambda - \lambda\beta)z} \\
&= \lambda\beta e^{-\lambda\beta z}.
\end{aligned}$$

Entonces se tiene que $S_N = Y_1 + \cdots + Y_N$ se distribuye como una exponencial con parámetro $\lambda\beta$.

En el caso de una distribución discreta para el monto o los valores de las variables aleatorias Y_i's, el cálculo de la convolución aleatoria es muy simple.

Ejemplo 2.5.2

Considere una cartera de seguros que puede producir 0, 1, 2, ó 3 reclamaciones en un período de tiempo fijo con probabilidades 0.1, .03, 0.4 y 0.2 respectivamente. Una reclamación individual podría tener un monto de 1,2 ó 3 pesos con probabilidad 0.5, 0.4 y 0.1 respectivamente. Calcular la función de densidad y distribución de la variable aleatoria de S_N, llamada en este caso la reclamación agregada.

Solución.

Se sabe por (2.5.30) que

$$
\begin{aligned}
f(S_N) &= \sum_{n=1}^{\infty} f^{(n)}(s)P_N(n) \\
&= \sum_{n=1}^{\infty} P(Y_1+Y_2+\cdots+Y_n=s)P(N=n) \\
&= \sum_{n=1}^{\infty}\left(\sum_{y\le s} P(Y_n=y)P(Y_1+Y_2+\cdots Y_{n-1}=s-y)\right)P(N=n) \\
&= \sum_{n=1}^{\infty}\left(\sum_{y\le s} P_Y(y)P^{(n-1)}(s-y)\right)P(N=n).
\end{aligned}
$$

De aquí que, de manera recursiva se obtienen los cálculos que se resumen en la siguiente tabla

Tabla 2.5.1						
(1)	(2)	(3)	(4)	(5)	(6)	(7)
X	$p^0(x)$	$p(x)$	$p^{(2)}(x)$	$p^{(3)}(x)$	f_{S_N}	F_{S_N}
0	1.0				0.10	0.1
1		0.5			0.15	0.25
2		0.4	0.25		0.22	0.47
3		0.1	0.40	0.125	0.215	0.685
4			0.26	0.300	0.164	0.849
5			0.08	0.315	0.095	0.944
6			0.01	0.184	0.0408	0.9848
7				0.063	0.0126	0.9974
8				0.012	0.0024	0.9998
9				0.001	0.0002	1.0000
n	0	1	2	3		
$P(n)$	0.1	0.3	0.4	0.2		

Dado que hay a lo más 3 reclamaciones y cada una produce un monto de a lo más 3 pesos, se pueden limitar los cálculos a $x = 0, 1, 2, \ldots, 9$.

La columna (2) lista la función de densidad de una variable aleatoria degenerada con toda la masa de probabilidad en cero. La columna (3) lista la función de densidad de la reclamación individual, y las columnas (4) y (5) se obtienen de aplicar (2.5.41). de manera recursiva.

Algunos cálculos se presentan a continuación:

Columna (4)

$$
\begin{aligned}
p^{(2)}(0) &= 0 \\
p^{(2)}(1) &= 1 \\
p^{(2)}(2) &= (0.4)(0) + (0.5)(0.5) + 0 = 0.25 \\
p^{(2)}(3) &= (0.4)(0.5) + (0.5)(0.4) = 0.20 \\
p^{(2)}(4) &= (0.5)(0.1) + (0.4)(0.4) + (0.5)(0.1) = 0.26 \\
&\vdots
\end{aligned}
$$

Columna (5) a través de (3) y (4)

$$
\begin{aligned}
p^{(3)}(0) &= 0\\
p^{(3)}(1) &= 0\\
p^{(3)}(2) &= 0\\
p^{(3)}(3) &= (0.25)(0.5) = 0.125\\
p^{(3)}(4) &= (0.4)(0.5) + (0.25)(0.40) = 0.300\\
p^{(3)}(5) &= (0.26)(0.50) + (0.4)(0.4) + (0.25)(0.1) = 0.315\\
p^{(3)}(6) &= (0.08)(0.5) + (0.26)(0.4) + (0.4)(0.1) = 0.1840\\
p^{(3)}(7) &= 0.063\\
p^{(3)}(8) &= 0.012\\
p^{(3)}(9) &= 0.001
\end{aligned}
$$

Columna (6)

$$
\begin{aligned}
f_{S_N}(0) &= 0.1\\
f_{S_N}(1) &= (0.5)(0.3) = 0.15\\
f_{S_N}(2) &= 0.12 + 0.10 = 0.22\\
f_{S_N}(3) &= 0.03 + 0.16 + 0.025 = 0.215\\
f_{S_N}(4) &= 0.104 + 0.06 = 0.164\\
f_{S_N}(5) &= 0.032 + 0.063 = 0.095\\
f_{S_N}(6) &= 0.004 + 0.0368 = 0.0408\\
f_{S_N}(7) &= 0.0126\\
f_{S_N}(8) &= 0.0024\\
f_{S_N}(9) &= 0.0002
\end{aligned}
$$

Columna (7)

$$
\begin{aligned}
F_{S_N}(0) &= 0.1\\
F_{S_N}(0) &= 0.1 + 0.15 = 0.25\\
&\vdots\\
F_{S_N}(9) &= 1.0
\end{aligned}
$$

Con estos ejemplos queda definido el concepto de sumas aleatorias de variables aleatorias que será de importancia en el desarrollo de este trabajo.

2.6 La distribución Poisson Compuesta.

Si N es una variable aleatoria *Poisson* con función de densidad dada por

$$P(N=n)=\frac{\lambda^n e^{-\lambda}}{n!}, \quad n=0,1,2,\dots \text{ y } \lambda>0.$$

y S_N está definida como

$$S_N = Y_1 + Y_2 + \cdots + Y_n$$

entonces se dice que S_N tiene una ***distribución Poisson Compuesta***.
La distribución ***Poisson Compuesta*** tiene características muy atractivas que serán presentadas posteriormente. Suponer dicha distribución tendrá sentido cuando la variabilidad del número de datos con los que se trabaje no diste mucho de su valor promedio, ya que $E(N)=Var(N)=\lambda$.

Cuando la varianza del número de reclamaciones excede su media, una distribución *Poisson* para la v.a. N no es recomendable, en esta situación el uso de una distribución binomial negativa es preferible, recordando que la densidad binomial negativa está dada por,

$$P(N=n)=\binom{r+n-1}{n}p^r q^n, \; n=0,1,2,\dots$$

Esta distribución tiene dos parámetros: $r>0$; $0<p<1$, $q=1-p$.

Para esta distribución se tiene

$$\begin{aligned} M(N;t) &= \left(\frac{p}{1-qe^t}\right)^r \\ E(N) &= \frac{rq}{p} \end{aligned}$$

y

$$Var(N)=\frac{rq}{p^2}.$$

Cuando una distribución binomial negativa se elige para N, la distribución para S_N se dice que es binomial negativa compuesta. Utilizando las fórmulas (2.5.35) y (2.5.36), se tiene

$$\begin{aligned} E(S_N) &= \frac{rq}{p}E(X) \\ Var(S_N) &= \frac{rq}{p}Var(X) + \frac{rq^2}{p^2}[E(X)]^2 \\ M(S_N;t) &= \left[\frac{p}{1-qM(X;t)}\right]^r \end{aligned}$$

Una familia de distribuciones para el número de reclamaciones puede generarse suponiendo que el parámetro Λ de la distribución Poisson es una v.a. con función de densidad $u(\lambda)$, $\lambda > 0$, y que la distribución condicional de N, dado $\Lambda = \lambda$, es Poisson con parámetro λ. Hay diversas situaciones en las cuales será útil considerar esta forma de definir la distribución de N. Por ejemplo, considere una población de asegurados donde existen varias clases de asegurados dentro de la población que genera el número de reclamaciones de acuerdo a la distribución Poisson, pero el parámetro de la distribución puede ser diferente en cada clase. Si la frecuencia relativa de los valores de λ están dadas por $u(\lambda)$, se puede adoptar la idea anterior utilizando la ***Ley de Probabilidades Totales***.

$$\begin{aligned} P(N=n) &= \int_0^\infty P[N=n|\Lambda=\lambda]u(\lambda)d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda}\lambda^n}{n!}u(\lambda)d\lambda. \end{aligned}$$

utilizando propiedades de esperanza condicional

$$E(N) = E(E(N|\Lambda)) = E(\Lambda)$$

ya que

$$E[N|\Lambda=\lambda] = \lambda,$$

y entonces

$$\begin{aligned} E[N|\Lambda] &= E[N|\Lambda=\lambda] \circ \Lambda = \Lambda \\ Var(N) &= E(Var(N|\Lambda)) + Var(E(N|\Lambda)) \\ &= E(\Lambda) + V(\Lambda) \end{aligned}$$

y

$$\begin{aligned} M(N;t) &= E[\exp\{tN\}] \qquad (2.6.41)\\ &= E\left[E[\exp\{tN\}\,|\Lambda]\right]\\ &= E\left[\exp\left\{\Lambda(e^t-1)\right\}\right]\\ &= M\left(\Lambda; e^t-1\right). \end{aligned}$$

La igualdad, $E[\exp\{tN\}\,|\Lambda] = \exp\{\Lambda(e^t-1)\}$, se sigue de la hipótesis de que la distribución condicional de N dado Λ, es Poisson con parámetro Λ. Así , como en el caso binomial negativo, $E(N) < Var(N)$.

Ejemplo 2.6.1
Para una v.a. S_N con distribución Poisson compuesta se tiene que

$$\begin{aligned} E(S_N) &= \lambda E(Y) \qquad (2.6.42)\\ Var(S_N) &= \lambda E(Y^2)\\ M(S_N;t) &= \exp\{\lambda(M(Y;t)-1)\}. \end{aligned}$$

Solución.

(1)

$$E(S_N) \overset{por(\ 2.5.35)}{=} E(Y)E(N) = \lambda E(Y)$$

(2)

$$\begin{aligned} Var(S_N) &\overset{por(\ 2.5.36)}{=} E(N)Var(Y) + [E(Y)]^2 Var(N)\\ &= \lambda Var(Y) + [E(Y)]^2\lambda = \lambda(Var(Y) + [E(Y)]^2)\\ &= \lambda E(Y^2). \end{aligned}$$

La función generadora de momentos de una v.a. Poisson N es

$$M(N;t) = \exp\{\lambda(e^t-1)\},$$

de aquí que utilizando (2.5.39)

$$\begin{aligned} M(S_N;t) &= M_N(\log M(Y;t))\\ &= \exp\{\lambda(\exp\{\log M(Y;t)\} - 1)\}\\ &= \exp\{\lambda(M(Y;t)-1)\}. \end{aligned}$$

$\square$

Propiedades del Proceso Poisson Compuesto.

Teorema 2.6.1 *Si $S_1, S_2, \ldots, S_m$ son variables aleatorias independientes tal que S_i tiene una distribución Poisson compuesta con parámetros λ_i y distribución de monto reclamado $P_i(x)$, $i = 1, 2, \ldots, m$, entonces $S = S_1 + S_2 + \cdots, +S_m$ tiene una distribución Poisson Compuesta con*

$$\lambda = \sum_{i=1}^{m} \lambda_i \qquad (2.6.43)$$

y

$$P(S \leq x) = P(x) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} P_i(x) \qquad (2.6.44)$$

Demostración.

Sea $M(i;t)$ la función generadora de momentos de $P_i(x)$, utilizando (2.6.42), la función generadora de momentos de S_i es

$$M(S_i;t) = \exp\{\lambda_i[M(i;t) - 1]\}.$$

Dada la independencia de $S_1, S_2, \ldots S_m$, la función generadora de momentos de su suma es

$$M(S;t) = \prod_{i=1}^{m} M(S_i;t) = \exp\left\{\sum_{i=1}^{m} \lambda_i [M(i;t) - 1]\right\},$$

finalmente, reescribiendo el exponente

$$M(S;t) = \exp\left\{\lambda\left[\sum_{i=1}^{m} \frac{\lambda_i}{\lambda}(M(i;t) - 1)\right]\right\}$$

donde

$$\lambda = \sum_{i=1}^{m} \lambda_i,$$

y dado que esta es la función generadora de momentos de la distribución Poisson Compuesta, especificada por (2.6.43) y (2.6.44), queda demostrado el teorema

□

Este resultado tiene dos consecuencias importantes en la construcción de modelos de seguros. Primero si consideran m portafolios de seguros independientes y éstos se combinan, entonces la reclamación agregada de los m portafolios tiene una distribución Poisson compuesta. Segundo, se puede considerar un único portafolio por un período de n años. Aquí se supone la independencia entre las m reclamaciones anuales agregadas y que cada una de estas tiene una distribución Poisson Compuesta, no es necesario que las distribuciones de las reclamaciones anuales sean idénticas. Dicho lo anterior, se sigue del teorema que la distribución de la reclamación agregada para un período de n-años tendrá una distribución Poisson Compuesta.

Ejemplo 2.6.2

Sean $x_1, x_2, \ldots, x_m$, m números diferentes, y suponga que $N_1, N_2, \ldots, N_m$ son v.a. mutuamente independientes, además suponga que N_i, con $i = 1, 2, \ldots, m$ tiene una distribución Poisson con parámetro λ_i, determinar la distribución de

$$x_1 N_1 + x_2 N_2 + \cdots + x_m N_m. \tag{2.6.45}$$

Solución

$x_i.N_i$ tiene distribución Poisson compuesta con parámetros de la distribución Poisson λ_i y una distribución de monto reclamado degenerado en x_i, aplicando el ***Teorema 2.6.1*** se tiene que la suma descrita en (2.6.45) tiene una distribución compuesta con parámetro

$$\lambda = \sum_{i=1}^{m} \lambda_i,$$

y función de probabilidad del monto reclamado $P(x)$, donde

$$P(x) = \begin{cases} \frac{\lambda_i}{\lambda} & \text{si } x = x_i,\ i = 1, 2, \ldots, m \\ 0 & \text{e.o.c.} \end{cases}$$

En este ejemplo puede suponerse que existen m portafolios independientes, cada individuo del portafolio i puede reclamar sólo x_i pesos, (la distribución del monto reclamado es degenerada, motivo por el cual hay un sólo valor de monto), y el número de personas que reclaman en el portafolio i es la variable aleatoria N_i

□

Se puede verificar que para cada distribución Poisson Compuesta con distribución de monto reclamado discreta puede ser representada como una suma de la forma (2.6.45). Sean $x_1, x_2, \ldots, x_m$ los valores discretos para las reclamaciones individuales y sean

$$\pi_i = P(x_i) \text{ para } i = 1, 2, \ldots, m$$

sus respectivas probabilidades. Sea N_i el número de términos en $S = X_1 + X_2 + \cdots + X_N$ que son iguales a x_i. Entonces agrupando términos se observa que

$$S = x_1 N_1 + x_2 N_2 + \cdots + x_m N_m. \tag{2.6.46}$$

En general, la N_i's de (2.6.46) son v.a. dependientes, sin embargo en el caso especial de una distribución Poisson Compuesta para S, estas variables son independientes, como se mostrará posteriormente.

Antes de enunciar el teorema que probará la aceveración anterior, se citarán algunas propiedades de la distribución multinomial que se usarán en la prueba del mismo.

Distribución Multinomial

Para la distribución multinomial, cada uno de n ensayos independientes puede resultar en m diferentes posibles resultados. La probabilidad de que un resultado caiga en el resultado i se denota por π_i. Se denotará a la v.a. que cuenta el número de resultados en n ensayos por N_i. Entonces

$$1 = \sum_{i=1}^{m} \pi_i \quad n = \sum_{i=1}^{m} N_i$$

y la función de densidad conjunta de $N_1, N_2, \ldots, N_m$ está dada por

$$P[N_1 = n_1, N_2 = n_2, \ldots, N_m = n_m] = \frac{n!}{n_1!n_2! \cdots n_m!}\pi_1^{n_1}\pi_2^{n_2}\cdots\pi_m^{n_m}$$

utilizando la densidad se tiene que la función generadora de la distribución multinomial es

$$\begin{aligned}
E\left(\exp\left\{\sum_{i=1}^{m} t_i N_i\right\}\right) &= \sum_{n_1,n_2,\ldots,n_m} \exp\{t_1n_1 + t_2n_2 + \cdots + t_mn_m\} \times \\
&\times\ P(N_1 = n_1, N_2 = n_2, \ldots, N_m = n_m) \\
&= \sum_{n_1,n_2,\ldots,n_m} \exp\left\{\sum_{i=1}^{m} t_i n_i\right\} \frac{n!}{n_1!n_2!\cdots n_m!} \times \\
&\times\ \pi_1^{n_1}\pi_2^{n_2}\cdots\pi_m^{n_m} \\
&= \sum_{n_1,n_2,\ldots,n_m} \frac{n!}{n_1!n_2!\cdots n_m!} \times \\
&\quad \times (\pi_1 \exp\{t_1\})^{n_1} (\pi_2 \exp\{t_2\})^{n_2} \cdots (\pi_m \exp\{t_m\})^{n_m} \\
&= \sum_{n_1,n_2,\ldots,n_m} \frac{n!}{\prod_{i=1}^{m} n_i} \prod_{i=1}^{m} (\pi_i \exp\{t_i\})^{n_i} \\
&= \left(\sum_{i=1}^{m} \pi_i \exp\{t_i\}\right)^{n}.
\end{aligned}$$

Así,

$$E\left(\exp\left\{\sum_{i=1}^{m} t_i N_i\right\}\right) = (\pi_1 \exp\{t_1\} + \pi_2 \exp\{t_2\} + \cdots + \pi_m \exp\{t_m\})^n,$$

esto último ya que

$$\left(\sum_{j=1}^{k} a_j\right)^n = \sum \frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} a_i^{n_i}$$

donde la suma es sobre todos los enteros de $n_1, n_2, \ldots, n_k$ los cuales suman a n. Esta última expresión es conocida como *El Teorema Multinomial*[3]

[3] Ref. [6] Pág. 531

Teorema 2.6.2 *Si $S = x_1 N_1 + x_2 N_2 + \cdots + x_m N_m$, tiene una distribución Poisson Compuesta con parámetro λ y probabilidades discretas para los montos reclamados, dada por*

$$\pi_i = P(x_i), \text{ con } i = 1, 2, \ldots, m.$$

Entonces

1. *$N_1, N_2, \ldots, N_m$ son mutuamente independientes.*
2. *N_i tiene una distribución Poisson con parámetro $\lambda_i = \lambda\pi_i$, $i = 1, 2, \ldots, m$*

Demostración.

Note que para una número fijo de reclamaciones independientes (ensayos) donde cada reclamación puede ser cualquiera de los m tipos de cantidades que se pueden reclamar, el número de reclamaciones de cada cantidad es multinomial con parámetros $n, \pi_1, \pi_2, \ldots \pi_m$. De aquí que dado que

$$N = \sum_{i=1}^{m} N_i = n,$$

la distribución condicional de $N_1, N_2, \ldots, N_m$ es la distribución multinomial.

Para este caso

$$\begin{aligned} E\left[\exp\left\{\sum_{n=0}^{\infty} t_i N_i\right\}\right] &= \sum_{n=0}^{\infty} E\left[\exp\left\{\sum_{n=1}^{m} t_i N_i\right\} | N = n\right] P[N = n] \\ &= \sum_{n=0}^{\infty} (\pi_1 \exp\{t_1\} + \ldots \pi_m \exp\{t_m\})^n \frac{e^{-\lambda}\lambda^n}{n!} \\ &= \exp\{-\lambda\} \sum_{n=0}^{\infty} \left(\lambda \sum_{i=1}^{m} \pi_i e^{t_i}\right)^n \\ &= \prod_{i=1}^{m} \exp\left\{\lambda\pi_i\left(e^{t_i} - 1\right)\right\}. \end{aligned}$$

Dado que este el el producto de m funciones cada una dependientes de una sola variable t_i, esto muestra la independencia de las N_i's. Además, si se toma $t_i = t$ y $t_j = 0$ para $j \neq i$ se obtiene

$$E(\exp\{tN_i\}) = \exp\left\{\lambda\pi_i\left(e^t - 1\right)\right\},$$

la cual es la función generadora de momentos de una distribución Poisson con parámetro $\lambda\pi_i$ que prueba el punto (2) de este Teorema.

□

La fórmula $S = \sum_{i=1}^{m} x_i N_i$ y el *Teorema 2.6.2* proveen un método alternativo para calcular la distribución Poisson Compuesta con un monto de reclamación discreto.

Primero se calcula la distribución de $x_1N_1, x_2N_2, \ldots, x_mN_m$, pero esto es una tarea fácil, ya que tendrá muchas entradas en cero debido a que en el cálculo de la $P(x_iN_i = x)$ se deberá tener que $\frac{x}{x_i}$ debe ser entero. Luego se calcula la convolución de esas m distribuciones para obtener la distribución de S. Este método es muy conveniente cuando m es pequeño.

Aún cuando una distribución continua haya sido seleccionada para el monto de las reclamaciones, una aproximación discreta puede algunas veces usarse para aplicar este método alternativo y dar una aproximación satisfactoria para la distribución de S.

El siguiente ejemplo compara el método básico de convoluciones con el alternativo para realizar el cálculo de S.

Ejemplo 2.6.3

Suponga que S tiene una distribución Poisson con $\lambda = 0.8$ y los montos individuales de reclamación son 1, 2 ó 3 con probabilidades 0.25, 0.375 y 0.375 respectivamente. Calcular $f_s(x) = P[S = x]$ para $x = 1, \ldots, 6$

Solución.

1. Utilizando convoluciones, es decir, el método tradicional, visto en el *ejemplo 2.5.2* se tiene

Tabla 2.6.1								
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
X	p^{*0}	p^{*1}	p^{*2}	p^{*3}	p^{*4}	p^{*5}	p^{*6}	f_S
0	1							0.4493
1		0.250						0.0897
2		0.375	0.0625					0.1438
3		0.375	0.1875	0.0156				0.1624
4			0.3281	0.0703	0.00391			0.0499
5			0.2813	0.1758	0.02344	0.00098		0.4736
6			0.1406	0.2637	0.07617	0.00732	0.00024	0.0309
n	0	1	2	3	4	5	6	
*	0.4493	0.3595	0.1438	0.0383	0.0077	0.0012	0.00016	

donde (*) representa

$$\frac{e^{-.8}(0.8)^n}{n!}$$

2. Utilizando el método alternativo desarrollado en esta sección, se tiene la siguiente tabla

Tabla 2.6.2					
(1)	(2)	(3)	(4)	(5)	(6)
x	$P(N_1 = x)$	$P(2N_2 = x)$	$P(3N_3 = x)$	(2) * (3)	(4) * (5)
0	0.818731	0.740818	0.740818	0.606531	0.449329
1	0.163746			0.121306	0.089866
2	0.016375	0.222245		0.194090	0.143785
3	0.001092		0.222245	0.037201	0.162358
4	0.000055	0.033337		0.030974	0.049906
5	0.000002			0.005703	0.047360
6	0.000000	0.003334	0.033337	0.003288	0.030923
i	1	2	3		
λ_i	0.2	0.3	0.3		
	$\frac{e^{-.2}(.2)^x}{x!}$	$\frac{e^{-.3}(.3)^{x/2}}{(x/2)!}$	$\frac{e^{-.3}(.3)^{x/3}}{(x/3)!}$		

Para aplicar las fórmulas debe notarse que $m = 3$. $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $\lambda_1 = \lambda p(1) = 0.2$, $\lambda_2 = \lambda p(2) = 0.3$, $\lambda_3 = \lambda p(3) = 0.3$.

Primero se calculan las columnas (2),(3) y (4), las entradas no-cero son probabilidades Poisson en las que $P(N_i = \frac{x}{i})$ tiene sentido ya que $\frac{x}{i}$ es entero, para $i = 1, 2, 3, \ldots$. Luego se obtiene la convolución de las columnas (2) y (3) registrando el resultado en (5). Finalmente se calcula la convolución de las columnas (4) y (5) y se registra en la columna (6). Debe recordarse que no se han hecho todos los cálculos ya que sólo se desarrolló la distribución de S para $x = 1, 2, \ldots, 6$, pero

$$P(s \leq 6) = f(0) + f(1) + \cdots + f(6) = 0.97325$$

y esto no suma 1

□

La fórmula $\sum_{i=1}^{m} x_i N_i = S$ y el Teorema anterior tienen otra implicación. En

lugar de definir una distribución Poisson compuesta para S especificando el parámetro λ y la distribución discreta de los montos reclamados $P(x)$, es posible definirla en términos de las cantidades reclamadas $x_1, x_2, \ldots, x_m$ y los parámetros $\lambda_1, \lambda_2, \ldots \lambda_m$ de las distribuciones Poisson asociadas a N_i con parámetro λ_i. En términos de esta nueva definición de la distribución de S, se obtiene que

$$E(N_i) = Var(N_i) = \lambda_i,$$

y de la independencia de las N_i's se obtiene

$$E(S) = E\left(\sum_{i=1}^{m} x_i N_i\right) = \sum_{i=1}^{m} x_i \lambda_i$$

y

$$Var(S) = Var\left(\sum_{i=1}^{m} x_i N_i\right) = \sum_{i=1}^{m} x_i^2 \lambda_i$$

Teorema 2.6.3 *1. Si S tiene una distribución Poisson Compuesta especificada por λ y $P(x)$, entonces la distribución de*

$$Z = \frac{S - \lambda E(X)}{\sqrt{\lambda E(X^2)}} \xrightarrow{\lambda \to \infty} N(0,1) \qquad (2.6.47)$$

2. Si S tiene una distribución Binomial negativa compuesta con parámetros r, p y $P(x)$, entonces

$$Z = \frac{S - r\left(\frac{q}{p}E(X)\right)}{\sqrt{r\left(\frac{q}{p}\right)E(X^2) + r\left(\frac{q^2}{p^2}\right)(E(X))^2}} \overset{r\to\infty}{\longrightarrow} N(0,1) \qquad (2.6.48)$$

Demostración. Se probará el número (1) mostrando que

$$\lim_{\lambda\to\infty} M(Z;t) = \exp\left\{\frac{t^2}{2}\right\}$$

El punto (2) puede probarse con la misma estrategia, pero la prueba involucra más pasos, por lo cual no se incluirá la demostración.

$$M(Z;t) = M\left(S;\frac{t}{\sqrt{\lambda E(X^2)}}\right)\exp\left\{\frac{-\lambda E(X)t}{\sqrt{\lambda E(X^2)}}\right\},$$

utilizando las propiedades de la función generadora de momentos en (2.6.47) y haciendo uso de la función generadora de momentos de una distribución Poisson Compuesta se tiene

$$M(Z;t) = \exp\left\{\lambda\left[M_X\left(\frac{t}{\sqrt{\lambda E(X^2)}}\right) - 1\right] - \frac{\lambda E(X)t}{\sqrt{\lambda E(X^2)}}\right\} \qquad (2.6.49)$$

y sustituyendo la expansión en serie de Taylor para

$$M\left(X;\frac{t}{\sqrt{\lambda E(X^2)}}\right)$$

con

$$t' = \frac{t}{\sqrt{\lambda E(X^2)}}$$

$$M(X;t) = 1 + \frac{E(X)}{1!}t' + \frac{E(X^2)}{2!}t'^2 + \cdots$$

en (2.6.49) se tiene

$$M(Z;t) = \exp\left\{\frac{1}{2}t^2 + \frac{1}{6}\frac{1}{\sqrt{\lambda}}\frac{E(X^3)}{(E(X)^2)^{3/2}}t^3 + \cdots\right\}$$

y si $\lambda \to \infty$, $M_Z(t)$ se aproxima a $e^{t^2/2}$, que es la función generadora de momentos de una distribución normal estandarizada.

La distribución normal es simétrica y consecuentemente su tercer momento central es cero, sin embargo la distribución de reclamación agregada es frecuentemente sesgada. La siguiente tabla muestra que los terceros momentos de S bajo una distribución Poisson Compuesta y una Binomial Negativa no son cero, por lo cual esta aproximación normal tiende a dar muy malas aproximaciones, de aquí la necesidad de mejorarlas.

Para completar la tabla siguiente se utilizaron las propiedades de cumulantes.

Tabla 2.6.3		
Paso	Poisson Compuesta	Bin. Neg. Compuesta
$M_S(t)$	$\exp\{\lambda(M_X(t)-1)\}$	$\left(\frac{p}{1-qM_X(t)}\right)^r$
$\ln M_S(t)$	$\lambda(M_X(t)-1)$	$r\ln p - r\ln(1 - qM_X(t))$
$\frac{d^3}{dt^3}\ln M_S(t)$	$\lambda M_X'''(t)$	$\frac{rqM_X'''(t)}{1-qM_X(t)} + \frac{3rq^2M_X'(t)M_X''(t)}{(1-qM_X(t))^2} + \frac{2rq^3M_X'(t)^3}{(1-qM_X(t))^3}$
$\frac{d^3}{dt^3}\ln M_S(t)\mid_{t=0}$	$\lambda E(X^3)$	$\frac{rqE(X^3)}{p} + \frac{3rq^2E(X)E(X^2)}{p^2} + \frac{2rq^3(E(X))^3}{p^3}$

Cabe señalar que para montos de reclamación positivos, $P(0) = 0$, y el tercer momento central de S es positivo en cada caso.

La aproximación aquí propuesta consiste en aproximar la distribución de la reclamación agregada S por un distribución gamma trasladada donde el parámetro α , β y x_0 se seleccionan igualando el primer, segundo y tercer momento central de S con los de una distribución gamma trasladada, de aquí que

$$E(S) = x_0 + \frac{\alpha}{\beta}$$

$$E\left[(S - E(S))^3\right] = \frac{2\alpha}{\beta^3}$$

$$Var(S) = \frac{\alpha}{\beta^2}$$

y de aquí se obtiene que

$$\beta = \frac{2Var(S)}{E\left[(S - E(S))^3\right]}$$

$$x_0 = E(S) - \frac{2(Var(S)^2)}{E[(S - E(S)^3)]}$$

$$\alpha = \frac{4(Var(s))^3}{E((S - E(S))^3)^2}$$

Para una distribución Poisson compuesta este procedimiento lleva a

$$\alpha = \frac{4\lambda(E(X^2))^3}{(E(X^3))^2} \tag{2.6.50}$$

$$\beta = \frac{2E(X^2)}{E(X^3)}$$

$$x_0 = \lambda E(X) - \frac{2\lambda(E(X))^2}{E(X^3)}$$

se puede mostrar que si $\alpha \to \infty$, $\beta \to \infty$ y $x_0 \to -\infty$

$$x_0 + \frac{\alpha}{\beta} = \mu \text{ constante}$$

$$\frac{\alpha}{\beta} = \sigma^2 \text{ constante}$$

Así se requiere entonces una aproximación más general a la distribución de reclamación agregada que acomode el sesgo que existe, y para ello se introducirá la parte medular de este trabajo: las expansiones en series de Edgeworth, expuestas en el siguiente capítulo.

Capítulo 3

Resultados de la teoría asintótica.

Dentro de la Estadística, existen algunos problemas distribucionales que tienen soluciones exactas aplicables, sin embargo también existen aquellos en los que no se encuentran tales soluciones, o en el caso de que existan. pueden llegar a ser tan complicadas que presentan inconvenientes para ser usadas de manera directa. Un ejemplo de esto, es la obtención de la distribución de sumas de variables aleatorias y sumas aleatorias de variables aleatorias, visto ya en el Capítulo 2. Para estos casos, lo más conveniente es abordarlos mediante la *Teoría Asintótica o de Grandes Muestras*, donde se puede conseguir derivar aproximaciones suponiendo cierta cantidad de información de la pobalción total.

En este capítulo se podrá mostrar que cuando n crece hasta infinito, la distribución correspondiente se aproxima a alguna otra distribución simple y estandarizada, en particular a una distribución normal asociada al *Teorema del Límite Central*. Se mostrará también que la distribución normal es el término principal de una serie asintótica frecuentemente en potencias de $n^{-\frac{1}{2}}$ o n^{-1}, dicho término puede proporcionar una aproximación numérica adecuada, y los términos restantes pueden ser usados para obtener más exactitud en la aproximación o para estimar las condiciones bajo las cuales la aproximación simple y estandarizada sea más útil.

Las expansiones para obtener las aproximaciones planteadas pueden llegar a tener rigurosas propiedades matemáticas en el caso de que $n \to \infty$, pero dado que en aplicaciones prácticas n es un número finito solo queda determinar que tan buena es la aproximación, por tal razón los resultados asintóticos que aquí se presentarán tendrán como objetivo enfatizar el modo de acelerar la convergencia sin necesidad de que n sea tan grande.

Se mencionarán entonces los principales teoremas relacionados con la convergencia de sumas de variables aleatorias para n tendiendo a infinito, profundizando en el entendimiento y utilidad del los mismos, concluyendo en el resultado que generalise y acelere la convergencia sin necesidad de que n sea tan grande, el cual será demostrado en su totalidad.

3.1 Leyes de los Grandes Números para el caso Binomial, Bernoulli y Borel

Teorema 3.1.1 *Sea $\{X_n; n \in \mathbb{N}\}$ una sucesión de v.a. binomiales tales que*

$$P(X_n|\theta) = Binomial(X_n|n,\theta). \tag{3.1.1}$$

Si $Y_n = \frac{1}{n}X_n$ para todo $n \in \mathbb{N}$, entonces

$$\text{para todo } \epsilon \in \mathbb{R} \lim_{n\to\infty} P(|Y_n - \theta| \geq \epsilon) = 0, \tag{3.1.2}$$

Demostración. Sea $A_n = \{w \in \Omega; |Y_n(w) - \theta| \geq \epsilon\}$, entonces utilizando la desigualdad de Tchebyshev y sabiendo que $E(X_n) = n\theta$ se tiene

$$\begin{aligned} P(A_n) &= P[w \in \Omega; |X_n(w) - n\theta| \geq n\epsilon] \\ &= P\left[w \in \Omega; (X_n(w) - n\theta)^2 \geq n^2\epsilon^2\right] \\ &\leq \frac{Var(X_n)}{n^2\epsilon^2} = \frac{\theta(1-\theta)}{n\epsilon^2} \text{ para todo } \epsilon \in \mathbb{R}^+, \end{aligned}$$

así $\lim_{n\to\infty} P(A_n) = 0$, es decir

$$\lim_{n\to\infty} P[|Y_n - \theta| \geq \epsilon] = 0$$

o bien $Y_n \xrightarrow{P} 0$.

Algo interesante que refleja este teorema, es el hecho de que el resultado $Y_n \xrightarrow{P} \theta$ puede ser utilizado para justificar la asignación frecuentista de probabilidades.

También debe notarse que el factor de estandarización de la v.a. X_n con el cual se obtiene la Y_n es muy grande, de tal forma que pulveriza a la variable aleatoria, degenerándola en el límite a una constante. La duda que surge de manera inmediata para este caso de v.a. Binomiales es si existe algún factor de estandarización que permita obtener en el límite una v.a. con determinada propiedad distribucional.

El poseer un límite con propiedades distribucionales, (que de alguna manera sea simple), permite el cálculo de los cuantiles y probabilidades que servirán en el caso $n \to \infty$ para aproximar el comportamiento distribucional de la sucesi'on.

En el Capítulo 1 se observó que el concepto de límite en probabilidad se puede alcanzar de diversas formas, las cuales involucran cierto grado de fuerza en el modo en que la convergencia es alcanzada. La convergencia obtenida en el Teorema anterior es un tanto débil en el sentido de que está referida a una convergencia estocástica o en probabilidad, vale entonces la pena explorar si para la misma sucesión puede obtenerse otro tipo de convergencia más fuerte.

Teorema 3.1.2 (Ley fuerte de los Grandes Números, Borel(1909).) *Sea $\{X_n; n \in \mathbb{N}\}$, una sucesión de variables aleatorias con distribución binomial y parámetro θ, si $Y_n = \frac{1}{n}X_n$ para todo $n \in \mathbb{N}$, entonces*

$$P(\lim_{n\to\infty} Y_n = \theta) = 1, \tag{3.1.3}$$

es decir $Y_n \xrightarrow{c.s.} \theta$

Para la demostración de este teorema será necesario enunciar el siguiente resultado:

Corolario 3.1.1 *Bajo las condiciones del teorema anterior se tiene que para todo $\varepsilon \in \mathbf{R}^+$*

$$\sum_{n=1}^{\infty} P[|Y_n - \theta| \geq \varepsilon] < \infty$$

Demostración.

Puede obtenerse que

$$E(X_n - \theta)^4 = n\theta(1-\theta)(3n\theta(1-\theta) - 6\theta(1-\theta) + 1) \text{ para todo } n \in \mathbf{N},$$

y de igual manera utilizando la desigualdad de Tchebyshev

$$P[|Y_n - \theta| \geq \varepsilon] = P[(X_n - \theta)^4 \geq n^4\varepsilon^4] \leq \frac{\theta(1-\theta)(3n\theta(1-\theta) - 6\theta(1-\theta) + 1)}{n^3\varepsilon^4}$$

para todo $n \in \mathbf{N}$, y entonces

$$P(|Y_n - \theta| \geq \varepsilon) \leq \frac{1}{n^2\varepsilon^4},$$

y como

$$\sum_{n=0}^{\infty} \frac{1}{n^p} < \infty \text{ para todo } p > 1,$$

se sigue que

$$\sum_{n=1}^{\infty} P(|Y_n - \theta| \geq \varepsilon) < \infty.$$

Con esto se procederá a demostrar el *Teorema 3.1.2*

Demostración.

Por el corolario anterior se tiene que

$$\sum_{n=1}^{\infty} P[|Y_n - \theta| \geq \varepsilon] < \infty; \text{ para todo } \varepsilon \in \mathbf{R}^+,$$

y utilizando el *Lema de Borel-Cantelli*, implica que

$$P[|Y_n - \theta| \geq \varepsilon \text{ i.v.}] = 0 \text{ para todo } \varepsilon \in \mathbf{R}^+$$

por lo que del *Lema (1.4.2)* se sigue que

$$P[\lim_{n\to\infty} Y_n = \theta] = 1$$

□

3.2 Leyes de los Grandes Números: Caso General.

En la sección anterior se demostró la ley débil y fuerte de los grandes números para el caso binomial. En esta sección se darán resultados semejantes pero para sucesiones más generales de variables aleatorias.

Definición 3.2.1 *Sea $\{X_n; n \in \mathbb{N}\}$ una sucesión de v.a. no necesariamente definidas sobre el mismo espacio tales que $E(X_n) < \infty$ para todo $n \in \mathbb{N}$. Se dice que $\{X_n; n \in \mathbb{N}\}$ tiene la propiedad débil de los grandes números si*

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0. \tag{3.2.4}$$

donde

$$S_n = \sum_{i=1}^{n} X_n.$$

Definición 3.2.2 *Sea $\{X_n; n \in \mathbb{N}\}$ una sucesión de v.a. no necesariamente definidas sobre el mismo espacio tales que $E(X_n) < \infty$ para todo $n \in \mathbb{N}$. Se dice que $\{X_n; n \in \mathbb{N}\}$ tiene la propiedad fuerte de los grandes números si*

$$\frac{S_n - E(S_n)}{n} \xrightarrow{c.s.} 0. \tag{3.2.5}$$

donde

$$S_n = \sum_{i=1}^{n} X_n.$$

Si $\{X_n; n \in \mathbb{N}\}$ son v.a. independientes *Bernoulli*$(X; \theta)$, entonces por lo visto en la sección anterior $\frac{S_n}{n} \xrightarrow{c.s.} \theta$, o en forma equivalente

$$\frac{S_n - E(S_n)}{n} \xrightarrow{c.s.} 0,$$

lo que se mostrará en esta sección es que bajo ciertas condiciones, no es necesario que las v.a. tengan la misma distribución para garantizar la convergencia.

Los siguientes resultados son consecuencia inmediata de la desigualdad de Tchebyshev, vista en el capítulo 1.

Corolario 3.2.1 (Condición de Markov.) *Sea $\{X_n; n \in \mathbf{N}\}$ una sucesión de v.a.i. tal que $Var(X_n) = \sigma_n^2 < \infty$ para todo $n \in \mathbf{N}$. Si*

$$lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2 = 0 \qquad (3.2.6)$$

entonces $\{X_n; n \in \mathbf{N}\}$ tiene la propiedad débil de los grandes números.

Demostración.

Como $\{X_n; n \in \mathbf{N}\}$ son independientes tales que $Var(X_n) = \sum_{k=1}^{n}\sigma_k^2$, usando el corolario anterior y la desigualdad de Tchebyshev se tiene que para todo $\varepsilon \in \mathbf{R}^+$

$$P\left[\left|\frac{S_n - E(S_n)}{n}\right| \geq \varepsilon\right] = o(1)$$

satisfaciendo la definicion de *Ley fuerte de los grande números*

□

Corolario 3.2.2 *Bajo las mismas hipótesis del corolario anterior, si $\sigma_n^2 = \sigma^2$ para todo $n \in \mathbf{N}$ entonces $\{X_n; n \in \mathbf{N}\}$ cumple con la propiedad débil de los grandes números.*

Demostración.

Como

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\sigma^2 = \sigma^2\lim_{n\to\infty}\frac{1}{n} = 0$$

se tiene el resultado.

El corolario (3.2.1) lo que en esencia dice mediante la condición de Markov, es que la varianza de las v.a. en juego, no debe crecer en más de n^2 para que se pueda garantizar la convergencia en probabilidad.

La suposicón de que exista la varianza puede ser relajada, sólo que la demostración se dificulta, por lo cual se enunciarán los dos resultados más importantes al respecto, terminando con una conclusión práctica en cuanto al uso de los mismos en la Estadística.

Teorema 3.2.1 *Sea $\{X_n; n \in \mathbb{N}\}$ una sucesión de v.a.i. tales que*

$$E(X_n) < \infty \textit{ para todo } n \in \mathbb{N}$$

si

$$\sum_{n=1}^{\infty} Var(X_n) < \infty$$

entonces

$$\sum_{n=1}^{\infty} [X_n - E(X_n)]$$

converge casi seguramente.

Demostración. Suponga $E(X_n) = 0$ para todo n $\in \mathbb{N}$ y sea $S_n = \sum_{i=1}^{n} X_i$,

entonces S_n converge si y sólo si $S_j - S_k \to 0$, y esto sucede casi seguramente si y sólo si para todo $\varepsilon \in \mathbb{R}^+$

$$P\left[|S_n - S_{n+k}| \geq \varepsilon \text{ para todo } k \in \mathbb{N}\right] = o(1),\ n \to \infty,$$

o equivalentemente

$$P\left[|S_n - S_m| \geq \varepsilon \text{ para todo } m \geq n\right] = o(1),\ n \to \infty.$$

La expresión anterior es igual a

$$\begin{aligned}
& P\left[\bigcup_{m,n \geq n_\varepsilon} |S_n - S_m| \geq \varepsilon\right] = o(1)\ n_\varepsilon \to \infty \\
= \ & P\left[\bigcup_{k=n}^{\infty} |S_n - S_{n+k}| \geq \varepsilon\right] = o(1)\ ,\ n \to \infty \\
= \ & P\left[\bigcup_{k=1}^{\infty} |S_n - S_{n+k}| \geq \varepsilon\right] = o(1)\ ,\ n \to \infty.
\end{aligned}$$

Pero debido a que P es continua y la sucesión es monótona, y utilizando *La desigualdad de Kolmogorov* se tiene

$$\begin{aligned}
P\left[\bigcup_{k=1}^{n}|S_n - S_{n+k}| \geq \varepsilon\right] &= \lim_{n\to\infty} P\left[\bigcup_{k=1}^{\infty}|S_n - S_{n+k}| \geq \varepsilon\right] \\
&= \lim_{n\to\infty} P\left[\max_{k\leq n}|S_n - S_{n+k}| \geq \varepsilon\right] \\
&\leq \lim_{n\to\infty} \frac{Var(S_n - S_{n+k})}{\varepsilon^2} \\
&\leq \lim_{n\to\infty} \sum_{j=n+1}^{n+k} \frac{Var(X_j)}{\varepsilon^2} = 0
\end{aligned}$$

□

Corolario 3.2.3 (Condición de Kolmogorov) *Sea $\{X_n; n \in \mathbf{N}\}$ una sucesión de v.a. independientes con varianza finita t, sea $\{b_n; n \in \mathbf{N}\} \subset \mathbf{R}^+$ una sucesión creciente tal que $b_n \to \infty$, si*

$$\sum_{n=1}^{\infty} \frac{Var(X_n)}{b_n^2} < \infty,$$

entonces $\{X_n; n \in \mathbf{N}\}$ tiene la propiedad fuerte de los grandes números, esto es

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{c.s.} 0$$

Demostración.

Como

$$Var\left(\frac{X_n - E(X_n)}{b_n}\right) = \frac{Var(X_n)}{b_n^2} \text{ para todo } n \in \mathbf{N}$$

por la independencia y la condición de *Kolmogorov* se tiene que

$$\sum_{n=1}^{\infty} Var\left(\frac{X_n - E(X_n)}{b_n}\right) < \infty$$

y por el teorema anterior

$$\sum_{n=1}^{\infty} \frac{X_n - E(X_n)}{b_n}$$

converge casi seguramente. Pero

$$\frac{S_n - E(S_n)}{b_n} = \frac{1}{b_n} \sum_{k=1}^{n} b_k \left(\frac{X_k - E(X_k)}{b_k} \right),$$

entonces por el *Lema de Kronecker*,

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{c.s.} 0$$

Nota: El *Lema de Kronecker* es el siguiente, sean $\{b_n \in \mathbb{R}^+; n \in \mathbb{N}\}$ una sucesión creciente a infinito y $\{X_n; n \in \mathbb{N}\}$ tal que $\sum_{n=1}^{\infty} X_n = X < \infty$. Entonces

$$\frac{1}{b_n} \sum_{j=1}^{n} b_j X_j = o(1); \ n \to \infty$$

Si $b_n = n$ se tiene *la definición (3.2.2)*.

□

Es interesante puntualizar que en el caso general de variables aleatorias que no poseen distribución binomial, se llega al mismo resultado que para éstas, es decir, a una convergencia casi segura, pero que implica la anulación de cualquier propiedad distribucional en el límite, ya que se degenera la variable aleatoria en una constante, esto último debido al factor de estabilización de la varianza o de estandarización de la suma, la sucesión $\{b_n; n \in \mathbb{N}\}$, por lo cual es necesario explorar nuevos factores de estandarización de sumas que permitan lograr algún tipo de convergencia a una variable aleatoria no degenerada en el límite.

Es importante notar también que de este resultado se deriva toda una metodología en la estimación de parámetros en la *Teoría Estadística y de*

Simulación, ya que si $\{X_n; n \in \mathbf{N}\}$ es una sucesión de v.a. independientes y con la misma distribución, donde $E(X_n) = \mu$ y $Var(X_n) = \sigma^2$ para todo $n \in \mathbf{N}$ se tiene por el resultado anterior que

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{c.s.} E(\bar{X}).$$

Una aplicación muy útil de este hecho lo proporciona el siguiente ejemplo.

Ejemplo 3.2.1

Determinar la probabilidad de ganar en un juego de solitario.

(Por solitario se entenderá el juego de solitario estándar ejecutado con un paquete de 52 cartas y con una estrategia fija).

Un posible enfoque es comenzar con la hipótesis razonable de que todos los (52!) arreglos posibles del paquete de cartas tienen la misma probabilidad de ocurrencia y entonces el problema se reducirá a intetar determinar cuantos arreglos llevan a ganar un juego. Desafortunadamente no parece ser un método sistemático para determinar el número de arreglos que llevan a ganar, ya que (52!) es un número muy grande y la única manera de saber si un arreglo particular lleva a ganar es jugando, en cuyo caso el enfoque no funcionaría.

Al parecer la determinación de esta probabilidad es intratable matemáticamente, sin embargo, no todo está perdido, ya que la experimentación es una técnica útil.

Para el ejemplo suponga que la experimentación consiste en jugar un gran número de veces, o alternativamente se programa una computadora para hacerlo.

Después de jugar n veces, se toma

$$X_i = \begin{cases} 1 & \text{si en el i-ésimo resultado ganó} \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces X_i, para $i = 1, \ldots, n$ serán variables aleatorias bernoulli independientes para las cuales $E(X_i) = P$(ganar el solitario). Y utilizando la *Ley Fuerte de Grandes Números*, se sabe que

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\text{número de juegos ganados}}{\text{número de juegos jugados}},$$

tiende a $E(X_i)$ con probabilidad 1.

3.3 El Teorema de Límite Central, un resultado de convergencia débil.

El Teorema del límite central, tiene que ver con el comportamiento límite de sumas de variables aleatorias, así como las leyes de los grandes números. Por lo visto anteriormente se tiene que si

$$Y_n = \frac{\sum_{i=1}^{n} X_i}{n},$$

con $\{X_n; n \in \mathbb{N}\}$ variables aleatorias con esperanza finita, entonces $Y_n - E(Y_n) \to 0$. El teorema central del límite involucra sumas estandarizadas, esto es, si $E(X_n) = \mu_n$ y $Var(X_n) = \sigma_n^2 < \infty$ para todo $n \in \mathbb{N}$, entonces se define

$$T_n = \frac{\sum_{i=1}^{n} (Y_n - \mu)}{\sqrt{\sum_{i=1}^{n} \sigma_n^2}}, \tag{3.3.7}$$

y ahora el problema es encontrar condiciones sobre $\{X_n; n \in \mathbb{N}\}$, $\{\mu_n; n \in \mathbb{N}\}$ y $\{\sigma_n^2; n \in \mathbb{N}\}$ para que

$$T_n \xrightarrow{d} V, \tag{3.3.8}$$

donde V es una v.a. con distribución $N(V|0,1)$.

Existen diferentes condiciones sobre las sucesiones involucradas para que T_n cumpla (3.3.8), así hay diferentes maneras de mostrarlo. Una forma es usando funciones características y el Teorema de Continuidad, otra es discretamente de la definición de convergencia débil (Billingsley 1968). En este trabajo se tomará el el primer camino, y la condición más común que se presenta en los textos estadísticos acerca de la sucesiones T_n.

Teorema 3.3.1 (Teorema del Límite Central) *Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuídas con varianza finita y $S_n = \sum_{i=1}^{n} X_i$, entonces*

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] \xrightarrow{d} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\{-\frac{1}{2}y^2\}dy$$

Demostración.

Sean $Y_i = X_i - \mu$, para $i = 1, \ldots, n$ y $\varphi_{Y_i}(t)$ la función característica de Y_i, pero

$$\varphi_{Y_i}(t) = \varphi_Y(t) \text{ para todo } i = 1, \ldots, n \tag{3.3.9}$$

ya que las v.a. son independientes e idénticamente distribuidas.

Sea

$$Z_n = \left(\sigma\sqrt{n}\right)^{-1}\sum_{i=1}^{n}(X_i - \mu),$$

y

$$\begin{aligned}
\varphi_{Z_n}(t) &= E\left[\exp\left\{it\left(\sigma\sqrt{n}\right)^{-1}\sum_{i=1}^{n}(X_i - \mu)\right\}\right] \\
&= \varphi_{\sum_{i=1}^{n} Y_i}\left(t(\sigma\sqrt{n})^{-1}\right) \\
&= \prod_{i=1}^{n}\varphi_{Y_i}\left(t(\sigma\sqrt{n})^{-1}\right) \\
&= \left[\varphi_Y\left(t(\sigma\sqrt{n})^{-1}\right)\right]^n,\ n \geq 1
\end{aligned}$$

Es necesario para continuar utilizar la siguiente expansión: Si

$$\int_{\mathbb{R}} |X|^n dF_X$$

existe y es finita, entonces la función característica $\varphi_X(t)$ de F_X se puede escribir como

$$\varphi_X(t) = 1 + \sum_{k=1}^{n-1}\mu_k'\frac{(it)^k}{k!} + R_n(t),$$

donde μ'_k es el momento k alrededor de la media.

Utilizando este resultado

$$\varphi_Y(t) = 1 - \frac{1}{2}\sigma^2 t^2 + o(t); \; \frac{o(t)}{t^3} \to 0 \tag{3.3.10}$$

y de aquí que

$$\begin{aligned}
\left[\varphi_Y\left(t(\sigma\sqrt{n})^{-1}\right)\right]^n &= \left[1 - \frac{1}{2}\frac{\sigma^2 t^2}{n\sigma^2} + o\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \\
&= \left[1 - \frac{t^2/2}{n} + o\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n
\end{aligned}$$

y de aquí que

$$\begin{aligned}
\log\varphi_{Z_n}(t) &= n\log\varphi_Y(t(\sigma\sqrt{n})^(-1)) \\
&= n\log\left[1 - \frac{1}{2}\frac{t^2}{n} + o\left(\frac{t}{\sigma\sqrt{n}}\right)\right] \\
&= -\frac{t^2}{2} + no\left(\frac{t^2}{n}\right),
\end{aligned}$$

que implica

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = \exp\{-\frac{t^2}{2}\}$$

□

3.4 Comparación entre La Ley Fuerte de los Grandes números y el Teorema del Límite Central.

Como se observó en la sección anterior el teorema del límite central permite la convergencia a una variable aleatoria no degenerada, la cual además es muy simple, ya que su distribución es la *Normal estándar*. Para aclarar las difirencias entre la ley fuerte de los grandes números y el teorema del límite central se expondrá el siguiente análisis:

Suponga que se tiene una sucesión de variables aleatorias independientes y con la misma distribución $\{X_n; n \in \mathbf{N}\}$, tal que $E(X_n) = 0$ y $Var(X_n) = 1$ para todo $n \in \mathbf{N}$. Sea $\sum_{i=1}^{n} X_i$ la suma parcial hasta n. Se tienen entonces los siguientes resultados.

(1) Por el teorema central del límite:

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$$

(2) Por la ley fuerte de los grandes números

$$\frac{S_n}{b_n} \xrightarrow{c.s.} 0$$

donde $\{b_n; n \in \mathbf{N}\} \subset \mathbf{R}^+$ es una sucesión creciente, tal que $b_n \to \infty$ y

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{Var(X_i)}{b_n^2} < \infty$$

para este ejemplo se tendrá que

$$\sum_{i=1}^{\infty} \frac{1}{b_n^2} < \infty. \qquad (3.4.11)$$

dado que $Var(X_i) = 1$ para todo $i \in \mathbf{N}$
En este caso, sea $b_n = n^\alpha$ y α tal que la condición de Markov se cumpla, esto es,

$$\sum_{i=1}^{\infty} \frac{1}{n^{\alpha}} < \infty \text{ si y solo si } \alpha \geq 1$$

Así la ley fuerte quedará expresada como

$$\frac{S_n}{n^\alpha} \xrightarrow{c.s.} 0 \text{ si } \alpha \geq 1 \text{ y } n \in \mathbf{N}.$$

En particular si $\alpha = 1$

$$\frac{S_n}{n} \xrightarrow{c.s.} 0,$$

pero si $\alpha = \frac{1}{2}$

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1).$$

De aquí se observa que en los dos casos el factor de normalización n ó $\sqrt{n}$ son sucesiones crecientes, pero una crece más rápido que la otra y esto es exactamente lo que hace la diferencia en la manera en que convergen las dos sucesiones de variables aleatorias. En el primer caso y siempre que $\alpha > 1$ se tendrá que el factor de normalización n^{α} pulverizará a la sucesión amarrando el límite a una constante, pero si se desea convergencia a una variable aleatoria no degenerada es necesario elegir la sucesión normalizadora lentamente divergente, es decir que la constante α del factor n^{α} esté contenida en el intervalo $(0, 1)$ para que dicha sucesión no se dispare de manera abrupta.

En resumen se puede decir que el Teorema del Límite Central garantiza cierta velocidad adecuada en la sucesión para que la misma converja a un límite no degenerado, por lo que se puede considerar una mejor aproximación que la ley fuerte de los grandes números.

Ejemplo 3.4.1

En un seguro de vida a un año, la aseguradora promete pagar una cantidad de b unidades de dinero si el asegurado muere dentro de este año y no pagar nada si el asegurado sobrevive el año. La probabilidad de una reclamación en el año se denotará por q, así la variable aleatoria que denota el monto de la reclamación X, tiene una distribución dada por:

$$f(x) = P(X = x) = \begin{cases} 1-q & \text{si } x = 0 \\ q & \text{si } x = b \\ 0 & e.o.c. \end{cases}$$

Escribiendo la variable aleatoria X como: $X = Ib$, se tiene que

$$\begin{aligned} E(X) &= bq \\ E(X^2) &= b^2q \\ Var(X) &= b^2q(1-q), \end{aligned}$$

donde b es la constante de pago si el evento muerte ocurre e I es la variable aleatoria que vale 1 si la muerte ocurre y 0 en otro caso. Así $P(I = 0) =$

$1-q$ y $P(I=1)\equiv q$. Por lo cual I es una variable aleatoria Bernoulli con parámetro (q).

Bajo este enfoque individual suponga:

Una compañía de seguros emite contratos a un año por sumas aseguradas de 1 y 2 unidades de dinero a individuos con probabilidades de muerte de 0.02 ó 0.10. La siguiente tabla proporciona el número de individuos n_k en cada una de las cuatro clases creadas que tienen sumas aseguradas b_k y probabilidad de reclamación q_k.

Tabla 3.4.1

k	q_k	p^{b_k}	p^{n_k}
1	0.02	1	500
2	0.02	2	500
3	0.10	1	300
4	0.10	2	500

La compañía quiere recoger de esta población de 1800 asegurados, una cantidad en primas igual al cuantil del *95 por ciento* de la distribución del total de reclamaciones, para asegurar tener reserva suficiente para enfrentar cualquier eventualidad de muerte en la población. Más aún, la aseguradora quiere que cada porción individual de esta cantidad total sea proporcional a la reclamación individual esperada.

La porción individual j con media $E(X_j)$ será $(1+\theta)E(X_j)$, con $\theta>0$. Esta cantidad individual extra es el recargo de seguridad y θ es un recargo relativo de seguridad. Se desea calcular θ.

Solución

Sea $S=X_1+X_2+\cdots+X_{1800}$ el monto total reclamado en el año por la cartera de 1,800 clientes. Entonces el criterio para elegir θ es $P[S\leq$

$(1+\theta)E(S)] = 0.95$. Esta probabilidad es equivalente a

$$P\left[\frac{S-E(S)}{\sqrt{Var(S)}} \leq \frac{\theta E(S)}{\sqrt{Var(S)}}\right] = 0.95$$

ya que

$$\frac{(1+\theta)E(S)-E(S)}{\sqrt{S}} = \frac{\theta(E(S))}{\sqrt{Var(S)}}.$$

Si se supone que los 1,800 individuos son independientes y que por clases tienen la misma distribución se pude aplicar una aproximación utilizando el Teorema del Límite Central, para ello se calcula la media y la varianza en cada clase

Tabla 3.4.2					
k	q_k	b_k	$b_k q_k$	Varianza $b_k^2 q_k(1-q_k)$	n_k
1	0.02	1	0.02	0.0196	500
2	0.02	2	0.04	0.7084	500
3	0.10	1	0.10	0.0900	300
4	0.10	2	0.20	0.3600	500

así

$$E(S) = \sum_{j=1}^{1800} E(X_j) = \sum_{j=1}^{4} n_k\mu_k = 160$$

$$Var(S) = \sum_{j=1}^{1800} Var(X_j) = \sum_{j=1}^{4} n_k\sigma_k^2 = 256$$

y

$$\frac{\theta E(S)}{\sqrt{Var(S)}} = 1.645$$

por ser el cuantil de orden 0.95 en una distribución $N(0,1)$. De aquí que

$$\theta = 1.645\frac{\sqrt{Var(S)}}{E(s)} = 1.645\left(\frac{16}{160}\right) = 0.1645$$

Es decir que la Compañía Aseguradora debe recolectar de su cartera $E(S) = 160$ como prima más un recargo de $\theta E(S) = 0.1645 * 160 = 26.32$ como margen de seguridad.

Ejemplo 3.4.2

Suponga ahora un ejemplo similar en donde la cantidad reclamada o el monto de la reclamación es aleatorio, como en el caso del seguro de autos.

Los asegurados de una cartera de autos en una compañía aseguradora caen en dos clases.

Tabla 3.4.3

k	n_k	q_k	λ	L
1	500	0.10	1	2.5
2	2000	0.05	2	5.0

La distribución exponencial truncada se define como:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} \text{si } 0 \leq x < L \\ 1 \text{si } x \geq L \end{cases}$$

La cual es una mezcla con función de densidad $f(x) = \lambda e^{-\lambda x}$, $0 < x < L$ y una masa de probabilidad de $e^{-\lambda L}$ en el punto L.
Nuevamente, la probabilidad de que el total de las reclamaciones exceda la cantidad cobrada de prima de los asegurados debe ser de 0.05. Si se supone un recargo relativo de seguridad θ, que sea igual en las dos clases, calcular θ.

Solución:
Para resolver este problema nuevamente hay que definir la variable X que representa el monto individual. Pero X es una variable aleatoria que puede escribirse como $X = IB$ donde I es una variable aleatoria Bernoulli(q) que representa el evento de que ocurra o no un siniestro y B, la variable aleatoria que representa el monto aleatorio a reclamar si ocurre un siniestro.

Así calculando los momentos de X utilizando propiedades de esperanza condicional se tiene:

$$E(X) = E[E(X|I)]$$
$$Var(X) = Var[E(X|I)] + E[Var(X|I)],$$

escribiendo

$$\mu = E(B|I = 1)$$
$$\sigma^2 = Var(B|I = 1),$$

ahora se calcula primero

$$E(X|I = 0) = 0,$$

ya que si no hay siniestro no se paga nada, y

$$E(X|I = 1) = E(B|I = 1) = \mu,$$

así tomando la composición con la variable aleatoria I

$$E(X|I = i) \circ I = E(X|I))\mu I$$

De aquí que

$$\begin{aligned} E(X) &= E[E(X|I)] = E(\mu I) \\ &= \mu E(I) = \mu q. \end{aligned}$$

y

$$Var(E(X|I)) = Var(\mu I) = \mu^2 Var(I) = \mu^2 q(1-q).$$

También de manera análoga

$$\begin{aligned} Var(X|I = 0) &= 0 \\ Var(X|I = 1) &= Var(B|I = 1) = \sigma^2, \end{aligned}$$

entonces

$$Var(X|I) = Var(X|I = i) \circ I = \sigma^2 I,$$

por lo cual

$$E[Var(X|I)] = E(\sigma^2 I) = \sigma^2 E(I) = \sigma^2 q,$$

teniéndose

$$Var(X) = \mu^2 q(1-q) + \sigma^2 q.$$

Continuando entonces con el problema

$$\mu = E(B|I=1) = \int_0^l x\lambda e^{-\lambda x} + Le^{-\lambda L} = \frac{1-e^{-\lambda L}}{\lambda}$$

$$E(B^2|I=1) = \int_0^L x^2\lambda e^{-\lambda x}dx + L^2 e^{-\lambda L} = \frac{2}{\lambda^2}(1-e^{-\lambda L}) - \frac{2l}{\lambda}e^{-\lambda L}$$

$$\sigma^2 = E(B^2|I=1) - (E(B|I=1))^2 = \frac{1-2\lambda Le^{-\lambda L} - e^{-2\lambda L}}{\lambda^2},$$

usando los valores de los parámetros λ y L y aplicando las fórmulas $E(X)$ y $Var(X)$ se tiene

Tabla 3.4.4						
k	q_k	μ_k	σ_k^2	media	varianza	n_k
1	0.10	0.9179	0.5828	0.09179	0.13411	500
2	0.05	0.5000	0.2498	0.02500	0.02436	2000

Entonces

$$\begin{aligned} E(S) &= 500(0.9179) + 2000(0.02500) = 95.89 \\ Var(S) &= 500(0.13411) + 2000(0.02436) = 115.78 \end{aligned}$$

$$P(S \leq (1+\theta)E(S)) = 0.95$$

Entonces aplicando el Teorema del límite central.

$$\frac{\theta E(S)}{\sqrt{Var(S)}} = 1.645$$

y

$$\theta = \frac{1.645\sqrt{115.78}}{95.89} = 0.1846$$

Entonces se cobra por prima a la cartera entera un total de 95.89 más un porcentaje de recargo de seguridad de 17.70.

Ejemplo 3.4.3

Una compañía de seguros cubre a 16,000 personas que poseen pólizas a un año por las cantidades que se muestran a continuación.

Datos	
B_k	n_k
10,000	8,000
20,000	3,500
30,000	2,500
50,000	1,500
100,000	500

La probabilidad de una reclamación q_k que por cada una de las 16,000 personas, se suponen mutuamente independientes y es 0.02. La compañía quiere fijar un límite de retención. Para cada vida, el límite de retención es la cantidad bajo la cual la compañía de seguros podrá retener la totalidad del riesgo y sobre la cual necesitará comprar reaseguro de otra compañía llamada reaseguradora. Como criterio de decisión la compañía desea minimizar la probabilidad de que el pago de reclamaciones retenidas más la cantidad que ésta paga por reaseguro exceda 8,250,000. El reaseguro está permitido a un costo de 0.025 por unidad de cobertura (0.125 porciento del monto esperado por unidad, 0.02). Se considera la cartera cerrada. Pólizas nuevas vendidas durante el año no entran en este proceso de decisión. Calcular el límite de retención que minimice la probabilidad de que las reclamaciones retenidas más el costo del reaseguro exceda 8,250,000.

Solución:

Primero se calculará todo en unidades de 10,000, como paso ilustrativo, sea S la cantidad de reclamaciones retenidas pagadas cuando el límite de retención es (2) $*$ (20,000). El portafolio de negocios que se retiene está dado por.

Datos	
B_k	n_k
1	8,000
2	8,000

Entonces

$$E(S) = \sum_{k=1}^{2} n_k b_k q_k = 8000(1)(0.02) + 8000(2)(0.02)$$

y

$$\begin{aligned} Var(S) &= \sum_{k=1}^{2} n_k b_k^2 q_k (1 - q_k) \\ &= 8000(1)(0.02)(0.98) + 8000(4)(0.02)(0.98) = 784 \end{aligned}$$

Adicionalmente a las reclamaciones retenidas S, hay un costo de reaseguro. La cobertura total en el plan es de

$$8000(1) + 3500(2) + 2500(3) + 1500(5) + 500(10)$$

La cantidad retenida es $8000(1) + 8000(2) = 24000$
así, la cantidad total cedida en reaseguro es

$$35,000 - 24,000 = 11,000$$

y el costo del reaseguro es

$$11000(0.025) = 275.$$

Así para un límite de retención de 2, las reclamaciones retenidas más el costo de reaseguro es de $S + 275$.
El criterio de decisión se basa en

$$\begin{aligned} P(S + 275 > 825) &= P(S > 550) \\ &= P\left(\frac{S - E(S)}{\sqrt{Var(S)}} > \frac{550 - E(S)}{\sqrt{Var(S)}}\right) \\ &= P\left(\frac{S - E(S)}{\sqrt{Var(S)}} 2.5\right), \end{aligned}$$

usando la distribución normal se tiene que esto es aproximadamente 0.0062. Para terminar este ejemplo habría que escoger difernetes límites de retención, calcular los costos de reaseguro y las probabilidades correspondientes y observar donde se alcanza el mínimo. Esto puede hacerse fácilmente en una hoja de cálculo de ***EXCEL***. Se recomienda probar con límites de retención entre [30.000 - 50,000].

3.5 Polinomios de Hermite-Chebyshev.

Es imporante conocer este concepto, ya que es de gran importancia para el desarrollo de ***Las Expansiones en Series de Edgeworth***, que serán tratadas posteriormente.

Sea

$$\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

la densidad de una normal con media igual a cero y varianza igual a uno, y sea

$$D^r = \frac{\partial}{\partial x^r}.$$

la r-ésima derivada con respecto a x para $r = 1, 2, \ldots$

Obteniendo la r-ésima derivada de $\alpha(x)$ para $r = 1, 2, \ldots$ se tiene que

$$\begin{aligned}
D\alpha(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\
&= \alpha(x)(-x) \\
D^2\alpha(x) &= x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}(x^2 - 1) \\
&= \alpha(x)(x^2 - 1) \\
D^3\alpha(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}(2x) + (x^2 - 1)\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}(-x) \\
&= \alpha(x)(3x - x^3)
\end{aligned}$$

$$\begin{aligned}
D^4\alpha(x) &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}(3-3x^2)+(3x-x^3)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}(-x)\\
&= \alpha(x)(3-6x^2+x^4)\\
D^5\alpha(x) &= \alpha(x)(-15x+10x^3-x^5)\\
D^6\alpha(x) &= \alpha(x)(-15+45x^2-15x^4+x^6)\\
&\vdots
\end{aligned}$$

De esta forma puede observarse que los resultados que se van obteniendo son equivalentes a:

$$\begin{aligned}
(-D)\alpha(x) &= x\alpha(x)\\
(-D)^2\alpha(x) &= (x^2-1)\alpha(x)\\
(-D)^3\alpha(x) &= (x^3-3x)\alpha(x)\\
(-D)^4\alpha(x) &= (3-6x^2+x^4)\alpha(x)\\
(-D)^5\alpha(x) &= (15x-10x^3+x^5)\alpha(x)\\
(-D)^6\alpha(x) &= (-15+45x^2-15^4+x^6)\alpha(x)\\
&\vdots
\end{aligned}$$

Entonces los polinomios de *Hermite-Chebyshev* se definen como

$$\begin{aligned}
(-D)\alpha(x) &= H_1(x)\alpha(x)\\
(-D)^2\alpha(x) &= H_2(x)\alpha(x)\\
(-D)^3\alpha(x) &= H_3(x)\alpha(x)\\
(-D)^4\alpha(x) &= H_4(x)\alpha(x)\\
&\vdots\\
(-D)^r\alpha(x) &= H_r(x)\alpha(x),
\end{aligned} \tag{3.5.12}$$

donde $H_r(x)$ es de grado r en x y el coeficiente x^r es único.

Ahora, si se toma $\alpha(x-t)$, se tiene que

$$\alpha(x-t) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(x-t)^2\right\}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(x^2 - 2xt + t^2\right)\right\} \\
&= \alpha(x) \exp\left\{xt - \frac{t^2}{2}\right\}.
\end{aligned}
$$

desarrollando el segundo término por series de Taylor con respecto a t alrededor de cero se tiene

$$
\begin{aligned}
f(t) &= \exp\left\{xt - \frac{t^2}{2}\right\} \\
&= f(0) + f^1(0) + \frac{f^2(0)t^2}{2!} + \cdots + \frac{f^n(0)t^n}{n!}
\end{aligned}
$$

donde

$$
\begin{aligned}
f(0) &= 1 \\
f^1(0) &= \left[\exp\left\{tx - \frac{t^2}{2}\right\}(x - t)\right]_{t=0} \\
&= x \\
f^2(0) &= \left[-\exp\left\{tx - \frac{t^2}{2}\right\} + (x - t)^2 \exp\left\{tx - \frac{t^2}{2}\right\}\right]_{t=0} \\
&= x^2 - 1 \\
f^3(0) &= \left[\exp\left\{tx - \frac{t^2}{2}\right\}(t - x) - 2(x - t)\exp\left\{tx - \frac{t^2}{2}\right\}\right]_{t=0} \\
&+ \left[(x - t)^3 \exp\left\{tx - \frac{t^2}{2}\right\}\right]_{t=0} \\
&= x^3 - 3x \\
&\vdots
\end{aligned}
$$

de esta forma

$$
f(t) = 1 + x\frac{t}{1!} + (x^2 - 1)\frac{t^2}{2!} + (x^3 - 3x)\frac{t^3}{3!} + \cdots
$$

por lo tanto

$$
\alpha(x - t) = \alpha(x) + x\alpha(x)\frac{t}{1!} + (x^2 - 1)\alpha(x)\frac{t^2}{2!} + (x^3 - 3x)\alpha(x)\frac{t^3}{3!} + \cdots
$$

$$
\begin{aligned}
&= \alpha(x) - D^1\alpha(x)\frac{t}{1!} + D^2\alpha(x)\frac{t^2}{2!} - D^3\alpha(x)\frac{t^3}{3!} + \cdots \\
&= \sum_{j=0}^{\infty}\frac{(-1)^j}{j!}t^j D^j\alpha(x) \\
&= \sum_{j=0}^{\infty} H_j(x)\alpha(x)\frac{t^j}{j!}
\end{aligned}
$$

esta última expresión es conocida como ***la función generadora de polinomios de Hermite-Chebyshev.***

Obsérvese que $H_r(x)$ es el coeficiente de $\frac{t^r}{r!}$, es decir, $H_r(x)$ es de la forma,

$$
H_r(x) = x^r - \frac{r^{[2]}}{2 - 1!}x^{r-2} + \frac{r^{[4]}}{2^2 - 2!}x^{r-4} - \frac{r^{[6]}}{2^3 - 2!}x^{r-6} + \cdots \qquad (3.5.13)
$$

De este modo se obtiene una forma más simple de generar los polinomios de Hermite-Chebyshev[1]

$$
\begin{aligned}
H_0 &= 1 \\
H_1 &= x \\
H_2 &= x^2 - 1 \\
H_3 &= x^3 - 3x \\
H_4 &= 3 - 6x^2 + x^4 \\
H_5 &= 15x - 10x^3 + x^5 \\
H_6 &= -15 + 45x^2 - 15x^4 + x^6 \\
&\vdots
\end{aligned}
$$

Una vez definidos estos polinomios, se procederá a introducir la parte medular de este trabajo: Las Expansiones en Series de Edgeworth.

3.6 Expansión en Series de Edgeworth.

Previo a esta sección, se han dado algunas ideas generales acerca de métodos de aproximación, ahora toca el turno a expansiones que conciernen es-

[1] Ref. [1] Pág. 169-170)

pecíficamente al comportamiento de sumas aleatorias de variables aleatorias, conocidas como *Expansiones en Series de Edgeworth*.

Sea Y cualquier v.a. tal que $E(Y) = \mu$ y $Var(Y) = \sigma^2$ y cuyos cumulantes k_r existen para $r \geq 3$.
Por (2.2.14) se conoce que $k_1 = \mu$ y $k_2 = \sigma^2$. Si se estandariza la v.a. Y se tiene que,

$$Z = \frac{Y-\mu}{\sigma} = \frac{Y}{\sigma} - \frac{\mu}{\sigma}.$$

Por definición la función generadora de cumulantes para Z es

$$\begin{aligned} K(Z;t) &= K\left(\frac{Y}{\sigma} - \frac{\mu}{\sigma}\right) \\ &\overset{(2.2.12)}{=} \frac{-\mu t}{\sigma} + K\left(\frac{Y}{\sigma}, t\right) \\ &\overset{(2.3.30)}{=} -\frac{\mu t}{\sigma} + K\left(Y; \frac{t}{\sigma}\right), \end{aligned}$$

y por definición de la función generadora de cumulantes se tiene

$$\begin{aligned} K(Z;t) &= -\frac{\mu t}{\sigma} + \frac{\mu t}{\sigma} + \frac{\sigma^2 t^2}{\sigma^2 2!} + \frac{k^3 t^3}{\sigma^3 3!} + \frac{k^4 t^4}{\sigma^4 4!} + \cdots \\ &= \frac{t^2}{2!} + \frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} \end{aligned}$$

Recordando que $\rho_r = O(n^{1-\frac{r}{2}})$ por (2.4.33), se pueden retener los términos hasta el orden $O(n^{-\frac{3}{2}})$, obteniendo

$$K(Z;t) = \frac{t^2}{2!} + \frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} + O(n^{-\frac{3}{2}}). \qquad (3.6.14)$$

Para obtener la función generadora de momentos de la variable estandarizada Z, basta con tomar la exponencial de la ecuación (3.6.14), es decir

$$\begin{aligned} M(Z;t) &= \exp\left\{\frac{t^2}{2!} + \frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} + O\left(n^{-\frac{3}{2}}\right)\right\} \\ &= \exp\left\{\frac{t^2}{2}\right\} \exp\left\{\frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} + O\left(n^{-\frac{3}{2}}\right)\right\}, \end{aligned}$$

expandiendo en series de Taylor la segunda exponencial alrededor del origen y conservando los términos hasta el orden $\mathcal{O}(n^{-\frac{3}{2}})$ se tiene,

$$M(Z;t) = \exp\left\{\frac{t^2}{2}\right\} \sum_{i=0}^{\infty} \frac{\left(\frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} + \mathcal{O}(n^{-3/2})\right)^i}{i!}$$

Por (2.4.33) se sabe que $\rho_3 = \mathcal{O}(n^{-\frac{1}{2}})$ y $\rho_4 = \mathcal{O}(n^{-1})$, lo que implica que si $i = 3, 4, 5, \ldots$, todos los términos serán de orden mayor o igual que $n^{-\frac{3}{2}}$, por lo tanto al dividir entre $n^{-\frac{3}{2}}$, los términos tienden a cero en el límite en lugar de acotarse, razón por la que pueden ser eliminados.

Si $i = 2$ sólo queda el término ρ_3^2, ya que este es de orden $\mathcal{O}(n^{-1})$.

Si $i = 1$ se tomarán todos los términos.

Si $i = 0$ el término es uno.

Entonces

$$\begin{aligned} M(Z;t) &= \exp\left\{\frac{t^2}{2}\right\}\left[1 + \left(\frac{\rho_3 t^3}{3!} + \frac{\rho_4 t^4}{4!} + \mathcal{O}(n^{-\frac{3}{2}})\right)\right] \\ &+ \left[\frac{(\frac{\rho_3 t^3}{6})^2 + \mathcal{O}(n^{-1})}{2!}\right] \\ &= \exp\left\{\frac{t^2}{2}\right\}\left[1 + \frac{\rho_3 t^3}{6} + \frac{\rho_4 t^4}{24} + \frac{\rho_3^2 t^6}{72}\right] + \mathcal{O}(n^{-\frac{3}{2}}) \end{aligned}$$

Esta función generadora de momentos debe ser invertida para encontrar la función de densidad, para lo cual se usará la siguiente expresión.

$$\int_{-\infty}^{+\infty} e^{ty}\phi(y)H_r(y)dy = t^r e^{\frac{t^2}{2}}, \tag{3.6.15}$$

Demostración.

Si $Y \sim N(0,1)$, entonces su función generadora de momentos es

$$M(X;t)=\int_{-\infty}^{\infty} e^{ty}\phi(y)dy=\exp\left\{\frac{t^2}{2}\right\}$$

si se multiplica t^r en ambos lados

$$t^r\int_{-\infty}^{\infty} e^{ty}\phi(y)dy=t^r\exp\left\{\frac{t^2}{2}\right\}$$

pero

$$\int_{-\infty}^{\infty} t^r e^{ty}\phi(y)dy=\int_{-\infty}^{\infty} e^{ty}H_r(y)\phi(y)dy$$

y esto se puede probar por inducción sobre k

para $k=1$

$$\begin{aligned}
\int_{-\infty}^{\infty} te^{ty}\phi(y)dy &= \int_{-\infty}^{\infty} t\frac{\partial e^{ty}}{\partial y}\phi(y)dy\\
&= \int_{-\infty}^{\infty} e^{ty}(-1)\frac{\partial\phi(y)}{\partial y}dy\\
&= \int_{-\infty}^{\infty} e^{ty}H_1(y)\phi(y)dy.
\end{aligned}$$

Ahora supóngase se cumple para $k=r$,

$$\int_{-\infty}^{\infty} t^r e^{ty}\phi(y)dy=\int_{-\infty}^{\infty} e^{ty}H_r(y)\phi(y)dy$$

entonces se demostrará para $k=r+1$,

$$\begin{aligned}
\int_{-\infty}^{\infty} t^{r+1}e^{ty}\phi(y)dy &= \int_{-\infty}^{\infty} te^{ty}H_r(y)\phi(y)dy\\
&= \int_{-\infty}^{\infty} e^{ty}(-1)\frac{\partial\left((-1)^r\frac{\partial^r\phi(y)}{\partial y^r}\right)}{\partial y}dy\\
&= \int_{-\infty}^{\infty} e^{ty}(-1)^{r+1}\frac{\partial^{r+1}\phi(y)}{\partial y^{r+1}}dy\\
&= \int_{-\infty}^{\infty} e^{ty}H_{r+1}(y)\phi(y)dy,
\end{aligned}$$

por lo tanto

$$\int_{-\infty}^{\infty} e^{ty} H_r(y)\phi(y)dy = t^r \exp\left\{\frac{t^2}{2}\right\}$$

□

Una vez demostrada la expresión (3.6.15), se procederá a invertir la función generadora de momentos, para lo cual se escribirá $M(Z:t)$ de la siguiente forma.

$$\begin{aligned} M(Z;t) &= t^0 \exp\left\{\frac{t^2}{2}\right\} + \frac{\rho_3 t^3}{6}\exp\left\{\frac{t^2}{2}\right\} + \\ &+ \frac{\rho_4 t^4}{24}\exp\left\{\frac{t^2}{2}\right\} + \frac{\rho_3^2 t^6}{72}\exp\left\{\frac{t^2}{2}\right\} + O(n^{-\frac{3}{2}}) \end{aligned} \quad (3.6.16)$$

por (3.6.15), se tiene que

$$\begin{aligned} M(Z;t) &= \int_{-\infty}^{+\infty} e^{ty}\phi(y)H_0(y)dy + \frac{\rho_3}{6}\int_{-\infty}^{+\infty} e^{ty}\phi(y)H_3(y)dy + \\ &+ \frac{\rho_4}{24}\int_{-\infty}^{+\infty} e^{ty}\phi(y)H_4(y)dy + \\ &+ \frac{\rho_3^2}{72}\int_{-\infty}^{+\infty} e^{ty}\phi(y)H_6(y)dy + O(n^{-\frac{3}{2}}) \\ &= \int_{-\infty}^{+\infty} e^{ty}\phi(y)\left[H_0(y) + \frac{\rho_3}{6}H_3(y) + +\frac{\rho_4}{24}H_4(y) + \frac{\rho_3^2}{72}H_6(y)\right]dy \\ &+ O(n^{-3/2}) \end{aligned}$$

lo que implica que

$$f_Z(y) = \phi(y)\left[1 + \frac{\rho_3}{6}H_3(y) + \frac{\rho_4}{24}H_4(y) + \frac{\rho_3^2}{72}H_6(y)\right] + O(n^{-\frac{3}{2}}). \quad (3.6.17)$$

De aquí que cualquier v.a. Y que posea todos sus momentos finitos y se estandarice, tiene una expansión en serie llamada de *Edgeworth*.

Ahora, por (2.3.29) y por la definición de cumulantes se tiene que,

$$\begin{aligned} K(S_n^*, t) &= -\frac{n\mu t}{\sigma\sqrt{n}} + n\left[\frac{\mu t}{\sigma\sqrt{n}} + \frac{t^2}{n2!} + \frac{k_3 t^3}{3!\sigma^3 n^{\frac{3}{2}}} + \frac{k_4 t^4}{4!\sigma^4 n^2} + \cdots\right], \\ &= \frac{t^2}{2} + \frac{k_3 t^3}{3!\sigma^3 n^{\frac{1}{2}}} + \frac{k_4 t^4}{4!\sigma^4 n} + \cdots + \frac{k_r t^r}{r!\sigma^r n^{1-\frac{r}{2}}} + \cdots. \end{aligned}$$

Recordando que

$$\rho_i = \frac{k_i}{\sigma^i} = \mathcal{O}(n^{1-i/2})$$

se tiene que

$$K(S_n^*;t) = \frac{t^2}{2} + \frac{\rho_3 t^3}{6n^{-1/2}} + \frac{\rho_4 t^4}{24n} + \mathcal{O}(n^{-\frac{3}{2}}). \qquad (3.6.18)$$

Debe observarse que (3.6.18) es idéntica a (3.6.17), de tal forma que siguiendo un procedimiento análogo para obtener $f_Z(y)$ se puede obtener $f_{S^*}(y)$, quedando de la siguiente forma,

$$f_{S_n^*}(y) = \phi(y)\left[1 + \frac{\rho_3 H_3(y)}{6\sqrt{n}} + \frac{\rho_4 H_4(y)}{24n} + \frac{\rho_3^2 H_6(y)}{72n}\right] + \mathcal{O}(n^{-\frac{3}{2}}), \quad (3.6.19)$$

integrando esta función se obtiene la distribución,

$$F_{S_n^*}(y) = \Phi(y) - \phi(y)\left[\frac{\rho_3 H_2(y)}{6\sqrt{n}} + \frac{\rho_4 H_3(y)}{24n} + \frac{\rho_3^2 H_5(y)}{72n}\right] + \mathcal{O}(n^{-\frac{3}{2}}). \quad (3.6.20)$$

3.7 Inversión de Cornish-Fisher.

Retomando la *Expansión en Series de Edgeworth*, para la función de distribución estandarizada S_n^*, defínase $k_{n\alpha}^*$ como el cuantil de orden α, es decir,

$$F(S_n^*, k_{n\alpha}^*) = \alpha. \qquad (3.7.21)$$

Por otro lado, se puede calcular otro cuantil k_α de la expresión $\Phi(k_\alpha) = \alpha$, donde k_α es el cuantil de una distribución *Normal Estándar*. Así utilizando (3.6.20) y (3.7.21) se tiene que

$$\begin{aligned}\Phi(k_\alpha) &= \Phi(k_{n\alpha}^*) - \phi(k_{n\alpha}^*)\left[\frac{\rho_3}{6\sqrt{n}}H_2(k_{n\alpha}^*) + \frac{\rho_4}{24n}H_3(k_{n\alpha}^*) + \frac{\rho_3^2}{72n}H_5(k_{n\alpha}^*)\right] \\ &+ \mathcal{O}(n^{-3/2})\end{aligned}$$

pasando todo al lado izquierdo se tiene

$$\begin{aligned}\Phi(k_\alpha) &- \Phi(k_{n\alpha}^*) - \phi(k_{n\alpha}^*)\left[\frac{\rho_3}{6\sqrt{n}}H_2(k_{n\alpha}^*) + \frac{\rho_4}{24n}H_3(k_{n\alpha}^*) + \frac{\rho_3^2}{72n}H_5(k_{n\alpha}^*)\right] \\ &+ \mathcal{O}(n^{-3/2}) = 0\end{aligned}$$

Esta ecuación puede ser resuelta por el método de *Newton*, para el cual la solución $f(u) = 0$ está dada por[2]

$$u = \overline{u} - \frac{f(\overline{u})}{f'(\overline{u})} - \frac{1}{2}\frac{f''(\overline{u})}{f'(\overline{u})}\left[\frac{f(\overline{u})}{f'(\overline{u})}\right]^2 - \cdots$$

De esta forma, tomando $f(k^*_{n\alpha})$ como

$$\begin{aligned} f(k^*_{n\alpha}) &= \Phi(k_\alpha) - \Phi(k^*_{n\alpha}) - \phi(k^*_{n\alpha})\left[\frac{\rho_3}{6\sqrt{n}}H_2(k^*_{n\alpha}) + \right. \\ &+ \left. \frac{\rho_4}{24n}H_3(k^*_{n\alpha}) + \frac{\rho_3^2}{72n}H_5(k^*_{n\alpha})\right] + O(n^{-3/2}) \end{aligned}$$

Ahora, tomando $\ddot{u} = k_\alpha$ y $u = k^*_{n\alpha}$ y valuando la función f y sus derivadas en $\ddot{u} = k_\alpha$ se tiene

$$\begin{aligned} f(k_\alpha) &= \phi(k_\alpha)\left[\frac{\rho_3}{6\sqrt{n}}H_2(k_\alpha) + \frac{\rho_4}{24n}H_3(k_\alpha) + \right. \\ &\left. + \frac{\rho_3^2}{72n}H_5(k_\alpha)\right] + O(n^{-3/2}). \end{aligned} \tag{3.7.22}$$

Por otro lado, si se evalúa la primera derivada de (3.7.23) en k_α se obtiene

$$\begin{aligned} f'(k^*_{n\alpha})|_{k_\alpha} &= -\phi(k_\alpha)\left[\frac{\rho_3}{6\sqrt{n}}H_3(k_\alpha + \frac{\rho_4}{24n}H_4(k_\alpha) + \right. \\ &\left. \frac{\rho_3^2}{72n}H_6(k_{n\alpha})\right] \end{aligned} \tag{3.7.23}$$

Ahora, tomando la segunda derivada de (3.7.23) y evaluando nuevamente en k_α se tiene

$$f''(k^*_{n\alpha})|_{k_\alpha} = \phi(k_\alpha)\left[k_\alpha + O(n^{-1/2})\right]. \tag{3.7.24}$$

Mediante (3.7.23), (3.7.24) y (3.7.24) se puede calcular

$$k^*_{n\alpha} = k_\alpha - \frac{f(k_\alpha)}{f'(k_\alpha)} - \frac{1}{2}\frac{f''(k_\alpha)}{f'(k_\alpha)}\left[\frac{f(k_\alpha)}{f'(k_\alpha)}\right]^2 - \cdots$$

[2]Ref. [2] Pág. 109-110

de tal forma que

$$
\begin{aligned}
k_{n\alpha} &= k_\alpha + \left[\frac{\frac{\rho_3}{6\sqrt{n}}(k_\alpha^2-1)+\frac{\rho_4}{24n}(k_\alpha^3-3k_\alpha)+\frac{\rho_3^2}{72n}(k_\alpha^5-10k_\alpha^3+15k_\alpha)}{1+\frac{\rho_3}{6\sqrt{n}}(k_\alpha^3-3k_\alpha)+\mathcal{O}(n^{-1})}\right] \\
&+ \frac{1}{2}\left[\frac{k_\alpha+\mathcal{O}(n^{-1/2})}{1+\mathcal{O}(n^{-1/2})}\right]\left[\frac{\frac{\rho_3}{6\sqrt{n}}(k_\alpha^2-1)+\mathcal{O}(n^{-1})}{1+\mathcal{O}(n^{-1/2})}\right]^2 \\
&+ \mathcal{O}(n^{-3/2}).
\end{aligned}
$$

Finalmete después de algunas operaciones algebraicas se obtiene

$$k_{n\alpha}^* = \frac{\rho_3}{6\sqrt{n}}(k_\alpha^2-1)+\frac{\rho_4}{24n}(k_\alpha^3-3k_\alpha)-\frac{\rho_3^2}{36n}(2k_\alpha^3-5k_\alpha)+\mathcal{O}(n^{-3/2}). \quad (3.7.25)$$

Para probar la exactitud de la aproximación *Edgeworth* contra el *Teorema del Límite Central* se realizará el siguiente ejemplo.

Ejemplo 3.7.1

Sean $Y_1, Y_2, \ldots Y_n$ v.a.i. con distribución exponencial con parámetro $\lambda = 1$, considere la suma parcial hasta n

$$S_n = \sum_{i=1}^{n} Y_i.$$

donde $S_n \sim Gamma(n, 1)$.

Se sabe de la Teoría de Probabilidad que para una distribución exponencial cuya densidad es

$$f_Y(y) = \lambda \exp\{-\lambda y\}$$

su función generadora de momentos es

$$M(Y:t) = E(e^{tY}) = \frac{1}{1-\left(\frac{t}{\lambda}\right)},$$

y su función generadora de cumulantes es

$$k_Y(t) = log M(Y;t) = -log\left(1 - \frac{t}{\lambda}\right),$$

expandiendo el logaritmo en series de Taylor se tiene

$$\begin{aligned} k_1 &= \frac{1}{\lambda} = E(Y) \\ k_r &= (r-1)!\,\frac{1}{\lambda^r}\ \ (r > 1). \end{aligned}$$

Considere $n = 50$ y $\lambda = 1$, entonces $S_{50} \sim Gamma(50, 1)$. Si el problema radica en encotrar el cuantil de orden α en la distribución S_n, es decir

$$P(S_n \leq y) = \alpha,$$

entonces se puede usar tanto la aproximación del *Teorema del Límite Central* como la aproximación por *Edgeworth*. Estas dos aproximaciones pueden ser comparadas con el verdadero valor ya que se conoce la distribución S_{50}, por lo tanto se procede a lo siguiente,

1. Utilizando el *Teorema del Límite Central* se tiene

$$P(S_n \leq y) = P(S_n^* \leq y^*) = \Phi(y^*) = \alpha$$

con

$$y^* = \frac{y - nE(Y)}{\sqrt{n}\sigma_Y}$$

y $\Phi(\cdot)$ = la distribución normal estándar. Si se obtiene y^*, tal que $\Phi(y^*) = \alpha$ entonces se puede despejar a y en la estandarización, en este caso $n = 50$ y $E(Y) = \sigma_Y = 1$

2. Utilizando la *Inversión Cornish-Fisher* se tiene

$$P(S_n^* \leq k_{n\alpha}^*) = \alpha$$

y entonces

$$k_{n\alpha}^* = k_\alpha \frac{\rho_3}{6\sqrt{n}}(k_\alpha - 1)^2 + \frac{\rho_4}{24n}(k_\alpha^3 - 3k_\alpha) - \frac{\rho_3^2}{36n}(2k_\alpha^3 - 5k_\alpha)$$

con k_α tal que $\Phi(k_\alpha) = \alpha$, $n = 50$ y

$$\rho_3 = \frac{k_3}{\sigma^3} = 2$$
$$\rho_4 = \frac{k_4}{\sigma^4} = 6$$

y sin estandarizar el cuantil se obtiene

$$k_\alpha = k_\alpha^* \sqrt{n}\sigma_Y + nE(Y)$$

Como se observa en la tabla (3.7.1) y en la gráfica (3.7.1), la precisión es mejor (casi idéntica), cuando la aproximación se hace por *Cornish-Fisher* que por el *Teorema del Límite Central.* El único inconveniente que presenta este ejemplo es que en la práctica no se conoce la distribución de los términos de la suma, sólo se tienen muestras de los valores que toman las variables aleatorias.
Ejemplo 3.7.2

Se realizará el mismo ejemplo anterior, pero suponiendo que se cuenta con una muestra aleatoria de tamaño 50 simulada con el *Teorema de la Transformación Integral de Probabilidad*, para obtener valores de las variables aleatorias exponenciales independientes.

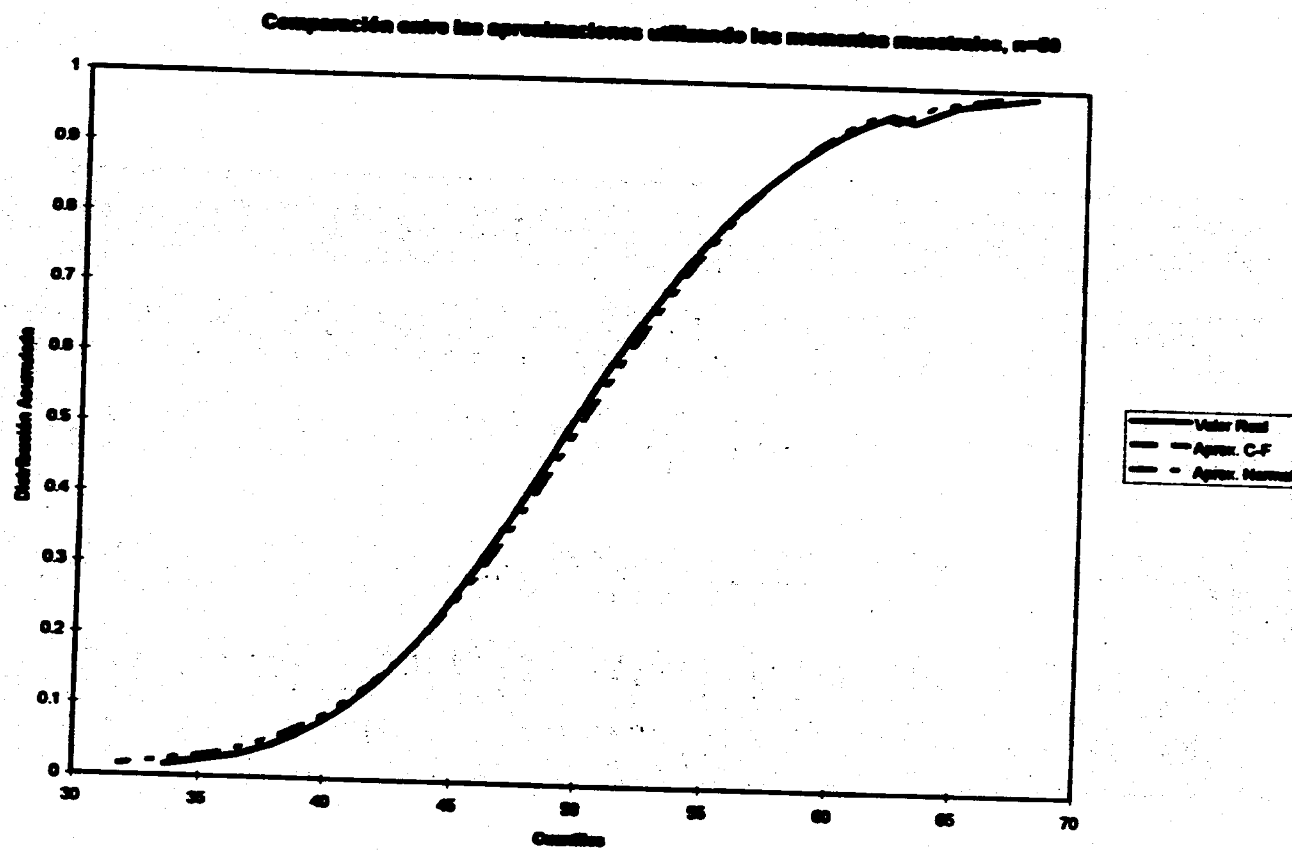
Como puede observarse en la Tabla (3.7.2) y en la gráfica (3.7.2), la aproximación normal muestral tiende a dar errores más altos en las colas que la aproximación Cornish-Fisher, resultando por este hecho más útil esta última, ya que en la práctica lo que se desea siempre, es encontrar los cuantiles muy chicos o muy altos, es decir, los que acumulan el 0.05 ó el 0.95.

3.8 Un ejemplo práctico.

El análisis sobre el ejemplo práctico se restringirá a períodos de un año, tales como cuando se evaluan los límites en los cuales los resultados de suscripción (emisión de pólizas), pueden presentar fluctuaciones que dependerán de

Comparación entre aproximaciones para el caso n=50						
Alpha	Valor Real	Dist.Norm.Inv.	Normal	Cornish-Fisher	Res.(C-F)	Res.(Norm.)
0.990	67.90331	2.32634	66.44972	67.90583	-0.00252	1.45359
0.975	64.78062	1.95996	63.85902	64.78185	-0.00123	0.92160
0.960	63.03970	1.75069	62.37922	63.04046	-0.00076	0.66048
0.950	62.17106	1.64485	61.63087	62.17163	-0.00057	0.54019
0.935	61.10877	1.51410	60.70633	61.10917	-0.00040	0.40243
0.920	60.23212	1.40507	59.93537	60.23238	-0.00026	0.29675
0.905	59.47910	1.31058	59.26720	59.47921	-0.00011	0.21190
0.890	58.81446	1.22653	58.67287	58.81451	-0.00004	0.14161
0.875	58.21634	1.15035	58.13420	58.21631	0.00002	0.08214
0.860	57.67004	1.08032	57.63902	57.66996	0.00008	0.03102
0.845	57.16525	1.01522	57.17870	57.16512	0.00014	-0.01345
0.830	56.69444	0.95416	56.74696	56.69428	0.00016	-0.05252
0.815	56.25196	0.89647	56.33902	56.25176	0.00020	-0.08707
0.800	55.83335	0.84162	55.95116	55.83314	0.00022	-0.11781
0.785	55.43518	0.78919	55.58043	55.43493	0.00025	-0.14525
0.770	55.05461	0.73885	55.22443	55.05433	0.00028	-0.16983
0.755	54.68832	0.69031	54.88122	54.68804	0.00028	-0.19190
0.740	54.33742	0.64334	54.54913	54.33712	0.00030	-0.21171
0.725	53.99726	0.59776	54.22681	53.99696	0.00030	-0.22952
0.710	53.66754	0.55338	53.91302	53.66721	0.00032	-0.24548
0.695	53.34698	0.51007	53.60677	53.34665	0.00033	-0.25979
0.680	53.03460	0.46770	53.30713	53.03425	0.00035	-0.27253
0.665	52.72947	0.42615	53.01332	52.72911	0.00036	-0.28385
0.650	52.43077	0.38532	52.72463	52.43042	0.00035	-0.29386
0.635	52.13781	0.34513	52.44041	52.13745	0.00036	-0.30260
0.620	51.84993	0.30548	52.16006	51.84956	0.00037	-0.31013
0.605	51.56653	0.26631	51.88311	51.56616	0.00037	-0.31658
0.590	51.28706	0.22754	51.60898	51.28670	0.00036	-0.32191
0.575	51.01107	0.18912	51.33727	51.01068	0.00038	-0.32620
0.560	50.73803	0.15097	51.06751	50.73764	0.00039	-0.32948
0.545	50.46752	0.11304	50.79931	50.46713	0.00039	-0.33179
0.530	50.19912	0.07527	50.53224	50.19872	0.00040	-0.33311
0.515	49.93243	0.03761	50.26589	49.93203	0.00040	-0.33346
0.500	49.66707	0.00000	50.00000	49.66667	0.00040	-0.33293
0.485	49.40264	-0.03761	49.73407	49.40225	0.00040	-0.33143
0.470	49.13878	-0.07527	49.46776	49.13838	0.00040	-0.32897
0.455	48.87513	-0.11304	49.20069	48.87472	0.00040	-0.32557
0.440	48.61129	-0.15097	48.93249	48.61089	0.00040	-0.32119
0.425	48.34689	-0.18912	48.66273	48.34650	0.00040	-0.31584
0.410	48.08155	-0.22754	48.39101	48.08115	0.00040	-0.30946
0.395	47.81486	-0.26631	48.11689	47.81446	0.00040	-0.30205
0.380	47.54636	-0.30548	47.83992	47.54596	0.00040	-0.29354
0.365	47.27558	-0.34513	47.55959	47.27520	0.00039	-0.28391
0.350	47.00229	-0.38532	47.27537	47.00190	0.00039	-0.27308
0.335	46.72567	-0.42615	46.98668	46.72529	0.00038	-0.26100
0.320	46.44520	-0.46770	46.69287	46.44481	0.00036	-0.24760
0.305	46.16049	-0.51007	46.39323	46.16013	0.00037	-0.23274
0.290	45.87063	-0.55338	46.08698	45.87028	0.00035	-0.21635
0.275	45.57491	-0.59776	45.77319	45.57456	0.00034	-0.19829
0.260	45.27246	-0.64334	45.45087	45.27214	0.00032	-0.17840
0.245	44.96229	-0.69031	45.11878	44.96198	0.00031	-0.15648
0.230	44.64322	-0.73885	44.77557	44.64293	0.00029	-0.13235
0.215	44.31390	-0.78919	44.41957	44.31362	0.00027	-0.10568
0.200	43.97267	-0.84162	44.04884	43.97242	0.00025	-0.07617

Tabla 3.7.1


Comparación entre las aproximaciones utilizando los momentos muestrales
Distribución Acumulada
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
30
35
40
45
50
55
60
65
70
Cuantiles
Valor Real
Aprox. C-F
Aprox. Normal

Gráfica 3.7.1

Comparación entre aproximaciones usando los momentos muestrales para n=50					
Alpha	Valor Real	Cornish-Fisher Real	Cornish-Fisher muestral	Normal Real	Normal muestral
0.990	67.6033	67.6068	61.7595	66.4497	61.0795
0.975	64.7806	64.7818	59.3628	63.8590	58.9176
0.960	62.1711	62.1716	57.3303	61.6309	57.0582
0.980	63.0397	63.0405	58.0099	62.3792	57.6827
0.945	61.7906	61.7911	57.0317	61.3009	56.7828
0.930	60.7998	60.8001	56.2512	60.4354	56.0805
0.915	59.9696	59.9697	55.5940	59.7030	55.4493
0.900	59.2480	59.2481	55.0213	59.0619	54.9143
0.885	58.6085	58.6086	54.5104	58.4878	54.4362
0.870	58.0281	58.0280	54.0467	57.9648	53.9968
0.855	57.4976	57.4975	53.6201	57.4821	53.5969
0.840	57.0049	57.0047	53.2236	57.0319	53.2202
0.825	56.5440	56.5439	52.8517	56.6086	52.8670
0.810	56.1089	56.1086	52.5006	56.2077	52.5324
0.795	55.6965	55.6963	52.1670	55.8256	52.2137
0.780	55.3065	55.3062	51.8484	55.4602	51.9087
0.765	54.9312	54.9310	51.5429	55.1087	51.6153
0.750	54.5706	54.5703	51.2487	54.7694	51.3321
0.735	54.2229	54.2226	50.9644	54.4407	51.0578
0.720	53.8863	53.8860	50.6888	54.1213	50.7913
0.705	53.5597	53.5594	50.4208	53.8101	50.5316
0.690	53.2420	53.2417	50.1597	53.5062	50.2780
0.675	52.9321	52.9318	49.9046	53.2088	50.0298
0.660	52.6292	52.6289	49.6547	52.9169	49.7869
0.645	52.3325	52.3322	49.4096	52.6294	49.5483
0.630	52.0413	52.0410	49.1687	52.3468	49.3163
0.615	51.7550	51.7546	48.9314	52.0674	49.0773
0.600	51.4730	51.4726	48.6973	51.7914	48.8470
0.585	51.1947	51.1943	48.4659	51.5182	48.6190
0.570	50.9198	50.9194	48.2370	51.2472	48.3929
0.555	50.6476	50.6472	48.0101	50.9780	48.1691
0.540	50.3778	50.3774	47.7847	50.7102	47.9447
0.525	50.1100	50.1097	47.5606	50.4434	47.7209
0.510	49.8438	49.8435	47.3376	50.1773	47.4980
0.495	49.5788	49.5784	47.1154	49.9114	47.2751
0.480	49.3146	49.3142	46.8934	49.6454	47.0521
0.465	49.0508	49.0505	46.6715	49.3789	46.8280
0.450	48.7872	48.7868	46.4493	49.1114	46.6105
0.435	48.5232	48.5228	46.2265	48.8426	46.3893
0.420	48.2586	48.2582	46.0026	48.5724	46.1667
0.405	47.9929	47.9924	45.7776	48.2998	45.9383
0.390	47.7258	47.7252	45.5513	48.0249	45.7036
0.375	47.4564	47.4560	45.3227	47.7469	45.4718
0.360	47.1849	47.1845	45.0918	47.4653	45.2369
0.345	46.9106	46.9101	44.8581	47.1797	44.9984
0.330	46.6327	46.6323	44.6211	46.8893	44.7561
0.315	46.3509	46.3505	44.3804	46.5937	44.5094
0.300	46.0646	46.0641	44.1363	46.2919	44.2576
0.285	45.7727	45.7724	43.8862	45.9833	44.0000
0.270	45.4749	45.4745	43.6294	45.6669	43.7369
0.255	45.1700	45.1697	43.3672	45.3413	43.4643
0.240	44.8570	44.8567	43.0975	45.0057	43.1842
0.225	44.5347	44.5344	42.8192	44.6584	42.8944
0.210	44.2016	44.2013	42.5311	44.2977	42.5934
0.195	43.8560	43.8557	42.2316	43.9216	42.2796

Table 3.7.2

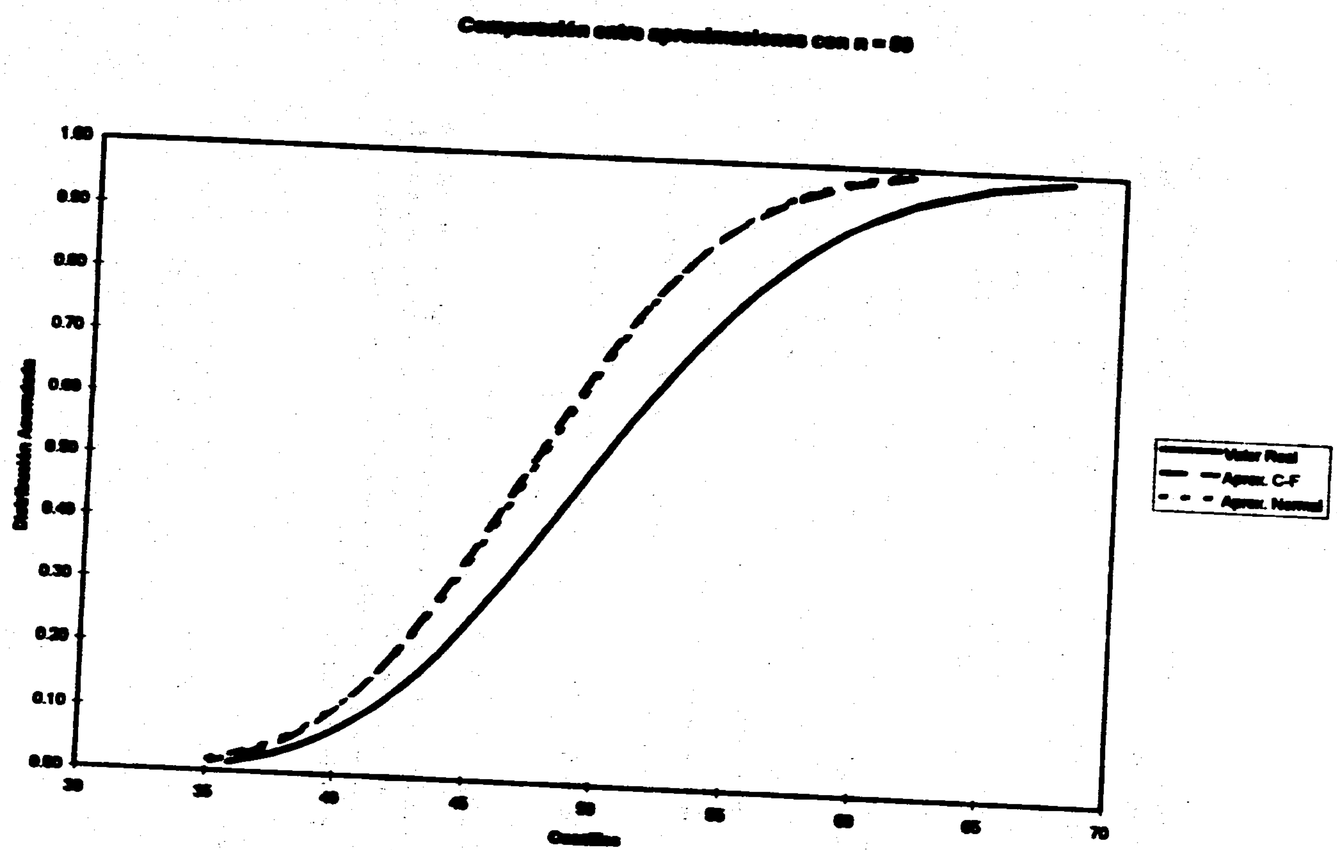

1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
30
35
40
45
50
55
60
65
70
Valor Real
Aprox. C-F
Aprox. Normal

figura 3.7.2

factores como lo son: el tamaño de la cartera, la distribución del tamaño de las reclamaciones, el reaseguro, niveles de cargo de seguridad, etc. Es claro que un año no es suficiente para considerar problemas importantes como la solvencia, pero si es suficiente para conocer el rango de fluctuación de las suscripciones anuales, es decir, saber si puede haber pérdida o ganancia. Para determinar estas fluctuaciones es necesario analizar el comportamiento de la reserva que se constituye para hacer frente a los riesgos que se generen en una cartera determinada, y para este análisis se harán los siguientes supuestos:

No se considerarán dividendos, productos finanacieros, etc., se dirá que toda reserva R está constituida por los siguientes elementos,

1. $(1+\theta)P$, que representa la prima total cobrada a los asegurados que constituyen la cartera, P = prima y θ = recargo de seguridad por desviaciones.

2. R_0 es la reserva inicial obtenida del cierre del año anterior.

3. S es el monto total de los siniestros reclamados del año en curso.

La relación de R con los elementos anteriores está determinada por

$$R = (1+\theta)P - R_0 - S$$

que refleja la cantidad de reserva al final del año.

De esta forma puede observarse que el problema fundamental para saber si existe ganancia o pérdida en un año es calcular la probabilidad de que R sea mayor que un límite inferior (conocido como margen de solvencia), y esta probabilidad sea alta, es decir

$$P(R \geq L_r) = \alpha \text{ , con } \alpha \to 1$$

pero

$$\begin{aligned} P(R \geq L_r) &= P[(1+\theta)P + R_0 - S \geq L_r] \\ &= P[S \leq (1+\theta)P + R_0 + L_R] \\ &= \alpha \end{aligned}$$

lo que significa que la probabilidad de que el monto de las reclamaciones sea menor a los ingresos de la aseguradora es alta.

Se observa que S es la parte que hace aleatoria a la variable R, ya que al inicio del año $(1+\theta)P$, R_0 y L_R son constantes conocidas. Es necesario señalar que el parámetro $(1+\theta)P$ no es realmente fijo, pues en la práctica la cartera de cualquier plan permanece abierta durante todo el año y por ende entran nuevos asegurados que pagan primas haciendo contratos con vigencia de un año rebasando el período de los contratos a un año emitidos con anterioridad. Por esta razón si se quiere suponer fijo dicho parámetro se tendrá que suponer que al inicio del año se tiene toda la cartera contratada. Bajo estos supuestos pueden hacerse análisis interesantes sobre S.

1. Dadas las constantes θ, P, R_0 y L_r, se quiere encontrar la probabilidad de que el monto por siniestros reclamados sea menor que el ingreso. Si esta probabilidad es alta entonces se tendrán síntomas de una buena selección de riesgos, es decir, se contará con una cartera sana.

2. Dado un nivel de probabilidad se quiere encontrar el respectivo cuantil que servirá para encontrar algunas de las constantes θ, P, R_0 y L_r.

3. También se puede determinar el límite de retención óptimo, como se mostró en el *ejemplo (3.4.3)*.

Con lo expuesto anteriormente se tienen las bases necesarias para introducir el siguiente ejemplo.

Se cuenta con una cartera correspondiente a us seguro de grupo en el ramo de vida con temporalidad de un año, lo que permite suponer independencia, con esta información se pretende conocer la distribución del monto reclamado por siniestros, ya que con ello se puede llegar a conocer $(1+\theta)P$ ó R_0 ó L_r.

En resumen, lo que se quiere determinar es

$$P(S \leq x) = \alpha \text{ donde } \alpha \to 1. \tag{3.8.26}$$

al conocer el valor que toma x, este puede ser expresado como

$$x = (1+\theta)P + R_0 - L_r.$$

Para calcular el valor x tal que cumpla (3.8.26), se propone usar la aproximación *Cornish-Fisher*, ya que con ella se podrá dar una solución aceptable al problema.

Suponga que se cuenta con una cartera de m pólizas en donde $S_1, S_2, \ldots S_m$ son las reclamaciones que se generan por póliza en un año, cada póliza tiene un número fijo de asegurados n_m, si se define

$$S = \frac{S_1}{n_1} + \frac{S_2}{n_2} + \cdots + \frac{S_m}{n_m} \quad (3.8.27)$$

como la suma de las reclamaciones promedio por póliza de toda la cartera entonces, la reclamación promedio anual puede ser escrita como

$$\frac{S}{m} = \frac{\frac{S_1}{n_1} + \frac{S_2}{n_2} + \cdots + \frac{S_m}{n_m}}{m}.$$

Ahora bien, por (3.8.26) se tiene que

$$P\left(\frac{S}{m} \leq x\right) = \alpha$$

y esto es lo mismo que

$$P(S \leq x') = \alpha \text{ con } x' = mx,$$

donde x representa la pérdida promedio anual para una aseguradora por concepto de reclamación.

Por facilidad para la expresión que determina el cuantil estandarizado por *Cornis-Fisher* (3.8.27) se escribirá como

$$S = S'_1 + S'_2 + \cdots + S'_m$$

donde

$$S'_i = \frac{S_i}{n_i}.$$

Para calcular x', se obtendrá primero el cuantil estandarizado $x^{*\prime}$ a través de la expresión (3.7.25), y recordando (2.2.22) y (2.2.23) se tiene.

$$\begin{aligned} x^{*\prime} &= k_\alpha + \left[\frac{\frac{\sum_{i=1}^{m}(S_i'-\bar{S}_i')^3}{m}}{s^3}\right]\frac{(k_\alpha^2-1)}{6\sqrt{m}} + \left[\frac{\frac{\sum_{i=1}^{m}(S_i'-\bar{S}_i')^4}{m}-3}{s^4}\right]\frac{(k_\alpha^3-3k_\alpha)}{24m} - \\ &- \left[\frac{\frac{\sum_{i=1}^{m}(S_i'-\bar{S}_i')^3}{m}}{s^3}\right]^2\frac{2k_\alpha-5k_\alpha}{36m} \end{aligned}$$

donde

$$\begin{aligned} \alpha &= \Phi(k_\alpha) \\ x' &= x^{*\prime}ds(S') + m(\bar{S}_i') \\ s &= \sqrt{\frac{\sum_{i=1}^{m}(S_i'-\bar{S}_i')}{m-1}} \end{aligned}$$

y finalmente

$$x = \frac{x'}{m},$$

cuyo valor es suficiente para dar solución a los puntos (1), (2) y (3), ya que la comoposición de R deja de ser estocástica.

Con este ejemplo sencillo y práctico se da por terminado este trabajo, ofreciendo las bases suficientes para continuar el investigando métodos de aproximación hasta campos más extensos.

Bibliografía.

[1] Kendall M.G. and Stuart. (1977)
"The Avanced Theory of Statistics"
Vol. 1
Mac. Millan New York.

[2] Beard, Pentikäinen and Pesonen. (1984)
"Risk Theory"
(The stochastics basis of insurance)
Chapman and Hall.

[3] Richard L. Burden y Douglas Faires. (1985)
"Análisis Numérico"
Grupo Editorial Iberoamericana.

[4] William Feller. (1986)
"Introducción a la Teoria de Probabilidad y sus Aplicaciones"
Vol. 1 y 2
Editorial Limusa.

[5] Bowers, Gerber, Hickman, Jones and Nesbitt. (1986)
"Actuarial Mathematics"
Society of Actuaries.

[6] Mood, Graybill and Boes. (1986)
"Introduction to the Theory of Statistics"
McGraw Hill.

[7] Ya-Lu Chou. (1988)
"Análisis Estadístico"
Interamericana.

[8] Bardorff-Nielsen and D.R. Cox. (1989)
"Asymptotic Techniques"
Chapman and Hall.