UNIVERSIDAD DE SONORA

División de Ciencias Exactas y Naturales

Departamento de Matemáticas

Desigualdades en la Teoría de Probabilidad

T E S I S

Que para obtener el título de:

Licenciado en Matemáticas

Presenta:

Azucena Campillo Navarro

Director de Tesis: Dr. Jesús Adolfo Minjárez Sosa

Hermosillo, Sonora, México, Agosto 2011

SINODALES

Dr. Jesús Adolfo Minjárez Sosa
Universidad de Sonora, Hermosillo, México.

Dr. Oscar Vega Amaya
Universidad de Sonora, Hermosillo, México.

Dr. José Arturo Montoya Laos
Universidad de Sonora, Hermosillo, México.

MC. Ma. Teresa Robles Alcaráz
Universidad de Sonora, Hermosillo, México.

Dedico este trabajo a mis padres, sé que el término de mi licenciatura representa principalmente el fruto de su trabajo.

Agradecimientos:

Agradezco profundamente a toda mi familia por su gran apoyo en todos los aspectos, en especial agradezco a mis padres por que recuerdo que desde muy pequeña me inculcaron el gusto y la necesidad de estudiar, les agradezco todo lo que soy y tengo como persona.

A mi director de tesis, Dr. Adolfo Minjárez, le agradezco todo su apoyo, enseñanza y colaboración para la elaboración de este trabajo, pero más aún le agradezco que haya sido mi profesor de todas las probabilidades que estudié, que fuera mi tutor, la confianza que me brindaba y que me haya regalado el gusto por la probabilidad con sus cátedras en clase.

Gracias a mis sinodales MC. Ma. Teresa Robles, Dr. Oscar Vega y Dr. José Arturo Montoya, todas sus observaciones, colaboraciones, consejos y tiempo que dedicaron para el estudio de mi tesis.

No puedo olvidar agradecer a una parte fundamental en mi formación como licenciada, mis compañeras de generación que durante los estudios se convirtieron en mis amigas y mi familia en la universidad, Dulce Yuridia, Marla y Carolina, les agradezco todo su apoyo, solo espero haberlas ayudado un poco de lo que ellas me ayudaron, y saben que estoy para ustedes siempre.

Por último, estoy muy agradecida con la mejor motivación que tuve durante todo el desarrollo de mi tesis, Fernando, le agradezco darme muchos ánimos, estar conmigo y ser más que mi mejor amigo.

Índice general

Introducción

Las desigualdades han sido un factor fundamental en el desarrollo de la matemática. Muchos teoremas importantes y famosos, que han sido el punto de partida de teorías matemáticas, están basados en alguna desigualdad clave. Por ejemplo, las desigualdades de Minkowski, Hölder y Schwarz en el análisis; las desigualdades de Fröbenius y de Schur para el rango y la traza de matrices, respectivamente, en el álgebra lineal; la desigualdad de Cauchy- Schwarz en varias áreas de la matemática, entre muchas otras.

El área de Probabilidad no ha sido la excepción. Por ejemplo, tenemos las desigualdades de Markov, Chebyshev, Bernstein y Lyapunov, las cuales han sido la base para demostrar varios teoremas límite como las Leyes de los Grandes Números y el Teorema del Límite Central, que han sido fundamentales en el desarrollo de la probabilidad y la estadística. Estas desigualdades han sido ampliamente estudiadas así como otras que son consecuencia o variantes de ellas, las cuales han tomado mayor importancia cuando se desea estimar una probabilidad o un parámetro de una distribución.

El objetivo del presente trabajo de tesis es estudiar y establecer desigualdades en la teoría de probabilidad fuera de las tradicionales. Para esto dividimos el trabajo en dos partes.

La Parte I consta de tres capítulos y corresponden al estudio de desigualdades de sumas de variables aleatorias independientes, para lo cual trabajamos principalmete con el libro [8] para el desarrollo de esta parte. En el Capítulo 1 presentamos desigualdades para el máximo de sumas de variables aleatorias independientes. Luego, en el Capítulo 2 establecemos cotas exponenciales para la probabilidad de sumas de variables aleatorias independientes. Por último, finalizamos esta parte con el Capítulo 3 estudiando desigualdades para momentos de sumas de variables aleatorias independientes.

Las desigualdades presentadas en la Parte I toman mayor importancia en el estudio de los teoremas límite donde se involucran las sumas de variables aleatorias independientes, por ejemplo en los teoremas de las Leyes de los Grandes Números y el Teorema del Límite Central. En este sentido, muchas de las desigualdades presentadas en este trabajo se pueden utilizar para refinar los resultados de convergencia.

La Parte II consta de dos capitulos. En el primero de ellos, que corresponde al Capítulo 4 del trabajo, presentamos las desigualdades tipo Bonferroni. Estas desigualdades surgen del problema de conteo al calcular la probabilidad del número de eventos que ocurren en una familia de ellos. En este caso podemos encontrar desigualdades en los dos sentidos, tanto inferior como superior. Para ilustrar la importancia de este tipo de desigualdades exponemos de manera general algunas posibles aplicaciones. Para el desarrollo de este capítulo nos hemos basado principalmente en el libro [3].

Finalizamos la Parte II presentando en el Capítulo 5 desigualdades para colas de distribuciones de variables aleatorias donde la cota superior esta determinada por la densidad de la variable aleatoria correspondiente. Esta propiedad observamos que, en principio, la satisface la distribución normal, sin embargo nuestro objetivo es extenderla a otras distribuciones.

Parte I

Desigualdades para Sumas de Variables Aleatorias Independientes

Capítulo 1

Desigualdades para el Máximo de Sumas

En este capítulo nos interesamos en estudiar desigualdades para la probabilidad del máximo de sumas de variables aleatorias independientes. Para ello, en la primera sección se establecen ciertas condiciones para los cuantiles de las sumas de variables aleatorias independientes. Mientras que, en la segunda sección, las desigualdades para la probabilidad del máximo de sumas dependen de condiciones sobre el primer y el segundo momento de cada sumando.

Para el desarrollo de este capítulo consideremos el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ donde están definidas las variables aleatorias independientes $X_1, X_2, \ldots, X_n$, y denotemos $S_k = \sum_{i=1}^{k} X_i$.

1.1 Condiciones sobre los Cuantiles

Definición 1.1. *Sea X una variable aleatoria y sea $0 < q < 1$. Un cuantil de orden q de una variable aleatoria X es cualquier número real, denotado por $k_q(X)$, que satisfaga las siguientes desigualdades*

$$P(X \leq k_q) \geq q, \quad P(X \geq k_q) \leq 1 - q.$$

Teorema 1.1. *Para cada valor $q \in (0,1)$ y para cada $x \in \mathbb{R}$*

$$P\left(\max_{1 \leq k \leq n} \{S_k - k_q(S_k - S_n)\} \geq x\right) \leq \frac{1}{q} P(S_n \geq x) \tag{1.1.1}$$

Demostración.

Denotemos los siguientes conjuntos

$$\bar{S}_k = \max_{1\leq l\leq k} \{S_l - k_q(S_l - S_n)\} \qquad (k = 1, 2, \ldots, n)$$

$$D_1 = [S_1 - k_q(S_1 - S_n) \geq x]$$

$$D_k = [\bar{S}_{k-1} < x, S_k - k_q(S_k - S_n) \geq x] \qquad (k = 2, 3, \ldots, n)$$

$$E_k = [S_n - S_k - k_{1-q}(S_n - S_k) \geq x] \qquad (k = 1, 2, \ldots, n)$$

Observemos que los eventos D_k y D_j, para $j \neq k$, no pueden ocurrir al mismo tiempo. En efecto, tomando $k < j$ y $j \leq n$, si el evento

$$D_k = [\bar{S}_{k-1} < x, S_k - k_q(S_k - S_n) \geq x]$$

ocurre, se sigue que $[S_k - k_q(S_k - S_n) \geq x]$ también sucede. Mientras que si el evento

$$D_j = [\bar{S}_{j-1} < x, S_j - k_q(S_j - S_n) \geq x]$$

ocurre, se sigue que $[S_k - k_q(S_k - S_n) < x]$ también ocurre, ya que $j > k$.
De esta manera se concluye que los eventos D_k y D_j, para $j \neq k$, no pueden ocurrir al mismo tiempo, y por lo tanto

$$P(D_k, D_j) = P(\emptyset) = 0,$$

donde $D_k, D_j = D_k \cap D_j$. Además, por lo anterior, ahora podemos poner

$$[\bar{S}_n \geq x] = \cup_{k=1}^{n} D_k,$$

y de aqui se obtiene

$$P[\bar{S}_n \geq x] = \sum_{k=1}^{n} P(D_k). \qquad (1.1.2)$$

Ahora mostraremos que

$$P(E_k) \geq q, \qquad para \quad k = 1, 2, \ldots, n. \qquad (1.1.3)$$

Para ello, primero observemos lo siguiente. Sea X una variable aleatoria, y sea q el orden de su cuantil, entonces

$$P[X \leq k_q(X)] \geq q \qquad y \qquad P[X \geq k_q(X)] \geq 1 - q,$$

de aquí se tiene que

$$P\left(-X \geq -k_q\left(X\right)\right) \geq q \qquad y \qquad P\left(-X \leq -k_q\left(X\right)\right) \geq 1-q,$$

por lo que $-k_q$ es un cuantil de orden $1-q$ para la variable aleatoria $-X$, es decir, $-k_q\left(X\right) = k_{1-q}\left(-X\right)$.

Aplicando estos argumentos en nuestro caso, podemos cambiar

$$k_{1-q}\left(S_n - S_k\right) = -k_q\left(S_k - S_n\right),$$

y, de esta manera, nos queda lo siguiente:

$$\begin{aligned}
E_k &= \left[S_n - S_k + k_{1-q}\left(S_n - S_k\right) \geq 0\right] \\
&= \left[S_n - S_k \geq -k_{1-q}\left(S_n - S_k\right)\right] \\
&= \left[S_n - S_k \geq k_q\left(S_k - S_n\right)\right].
\end{aligned}$$

Por la definición de cuantil podemos concluir que

$$P\left(E_k\right) \geq q.$$

Ahora, observermos que se cumple

$$\cup_{k=1}^{n}\left[D_k, E_k\right] \subset \left[S_n \geq x\right],$$

por medio de las siguientes relaciones:

$$\begin{aligned}
\left[D_k, E_k\right] &= \left[\bar{S}_{k-1} < x, S_k - k_q\left(S_k - S_n\right) \geq x, S_n \geq S_k + k_{1-q}\left(S_n - S_k\right)\right] \\
&= \left[\bar{S}_{k-1} < x, S_k \geq x + k_q\left(S_k - S_n\right), S_n \geq S_k - k_q\left(S_k - S_n\right)\right] \\
&= \left[\bar{S}_{k-1} < x, S_n \geq x + k_q\left(S_k - S_n\right) - k_q\left(S_k - S_n\right)\right] \\
&= \left[\bar{S}_{k-1} < x, S_n \geq x\right] \\
&\subset \left[S_n \geq x\right].
\end{aligned}$$

Además, podemos notar que los eventos $\left[D_k, E_k\right]$ y $\left[D_j, E_j\right]$, para $k \neq j$, son ajenos de la misma manera que los eventos D_k y D_j son ajenos, para $k \neq j$. Entonces

$$P\left(S_n \geq x\right) \geq P\left(\cup_{k=1}^{n} D_k, E_k\right) = \sum_{k=1}^{n} P\left(D_k, E_k\right).$$

La igualdad se debe a que los eventos son ajenos, y más aún, cabe notar que los eventos E_k y D_k son independientes ya que dependen de las variables aleatorias $X_i's$, las cuales por hipótesis son independientes. Aplicando esta última observación se obtiene

$$\sum_{k=1}^{n} P(D_k, E_k) = \sum_{k=1}^{n} P(D_k) P(E_k).$$

Para finalizar, retomemos las Desigualdades (1.1.2) y (1.1.3) con las cuales se obtiene lo siguiente:

$$\begin{aligned} P(S_n \geq x) &\geq \sum_{k=1}^{n} P(D_k) P(E_k) \\ &\geq q \cdot \sum_{k=1}^{n} P(D_k) \\ &= q \cdot P(\bar{S}_n \geq x). \end{aligned}$$

De esta manera queda demostrado el teorema. ■

Definición 1.2. *Una mediana mX, de una variable aleatoria X, es el cuantil $k_{\frac{1}{2}}(X)$.*

Teorema 1.2. *Desigualdades de Lévy.*
Para cada $x \in \mathbb{R}$ se cumple que:

$$P\left(\max_{1 \leq k \leq n} \{S_k - m(S_k - S_n)\} \geq x\right) \leq 2 \cdot P(S_n \geq x), \tag{1.1.4}$$

y para cada $x \geq 0$ tenemos que:

$$P\left(\max_{1 \leq k \leq n} |S_k - m(S_k - S_n)| \geq x\right) \leq 2 \cdot P(|S_n| \geq x). \tag{1.1.5}$$

Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes y simétricas (ver la definición en Apéndice (A)), entonces para cada $x \in \mathbb{R}$ se satisface la siguiente desigualdad:

$$P\left(\max_{1 \leq k \leq n} S_k \geq x\right) \leq 2 \cdot P(S_n \geq x), \tag{1.1.6}$$

y para cada $x \geq 0$ se cumple que:

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq x\right) \leq 2 \cdot P(|S_n| \geq x). \tag{1.1.7}$$

Demostración.

La demostración de la Desigualdad (1.1.4) se sigue inmediatamente del teorema (1.1), al sustituir $q = \frac{1}{2}$ en la Desigualdad (1.2.4).

Para demostrar la Desigualdad (1.1.5) notemos que el evento

$$\left[\max_{1\leq k\leq n} |S_k - m\,(S_k - S_n)| \geq x\right]$$

ocurre si uno de los siguientes dos eventos ocurre:

$$Caso \quad (i): \left[\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_n)\} \geq x\right].$$

$$Caso \quad (ii): \left[\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_n)\} \leq -x\right].$$

Para el caso (i), por la Desigualdad (1.1.4) tenemos que

$$P\left(\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_n)\} \geq x\right) \leq 2 \cdot P\,(S_n \geq x)$$

para $x \geq 0$.

Ahora, para el caso (ii), tenemos que

$$\left[\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_m)\} \leq -x\right] = \left[-\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_m)\} \geq x\right],$$

entonces,

$$P\left(\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_m)\} \leq -x\right) = P\left(-\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_m)\} \geq x\right)$$

y, por la Desigualdad (1.1.4) se obtiene que

$$\begin{aligned} P\left(-\max_{1\leq k\leq n} \{S_k - m\,(S_k - S_m)\} \geq x\right) &\leq 2 \cdot P\,(-S_n \geq x) \\ &= 2P\,(S_n \leq -x) \end{aligned}$$

para $x \geq 0$.

Para la demostración de la Desigualdad (1.1.6) debemos considerar los siguientes dos puntos:

- Toda variable aleatoria simétrica tiene media igual a cero.
- La suma de variables aleatorias simétricas es una variable aleatoria simetrizada (ver definición en Apéndice (A)).

Ahora, podemos observar la Desigualdad (1.1.6) se obtiene de la Desigualdad (1.1.4), ya que $m(S_k - S_n) = 0$.

Finalmente, la Desigualdad (1.1.7) se obtiene siguiendo la demostración de la Desigualdad (1.1.5) y tomando en cuenta la Desigualdad (1.1.6). ■

Teorema 1.3. *Si $P(S_n - S_k \geq -C) \geq q$, donde $k = 1, \ldots, n-1$, para alguna constante $C \geq 0$ y $q \geq 0$, entonces*

$$P\left(\max_{1\leq k\leq n} S_k \geq x\right) \leq \frac{1}{q} \cdot P(S_n \geq x - C) \tag{1.1.8}$$

para cada $x \in \mathbb{R}$.

Para la demostración del Teorema (1.3) ocuparemos el Lema (A.1) del Apéndice (A) y del siguiente lema.

Lema 1.1. *Sea C una constante no negativa y un conjunto de cuantiles:*

$$k_q(S_1 - S_n), \cdots, k_q(S_{n-1} - S_n)$$

de orden q, con $0 < q < 1$, tal que

$$k_q(S_k - S_n) \leq C, \quad para \quad k = 1, \cdots, n-1. \tag{1.1.9}$$

Entonces se cumple la Desigualdad (1.1.8) del Teorema (1.3).

Demostración.
Por la desigualdad (1.1.9) tenemos que:

$$[S_k - C \geq y] \subset [S_k - k_q(S_k - S_n) \geq y],$$

para cada $y \in \mathbb{R}$.
Así,

$$\left[\max_{1\leq k\leq n}\{S_k - C\} \geq y\right] \subset \left[\max_{1\leq k\leq n}\{S_k - k_q(S_k - S_n)\} \geq y\right],$$

y de aquí se obtiene que

$$P\left[\max_{1\leq k\leq n} S_k \geq y + C\right] \leq P\left[\max_{1\leq k\leq n} \{S_k - k_q(S_k - S_n)\} \geq y\right]$$

para $y \in \mathbb{R}$.

Por otro lado, recordemos que por el Teorema (1.1) tenemos que:

$$P\left[\max_{1\leq k\leq n} \{S_k - k_q(S_k - S_n)\}\right] \leq \frac{1}{q} P(S_n \geq y),$$

lo cual implica que

$$P\left[\max_{1\leq k\leq n} S_k \geq y + C\right] \leq \frac{1}{q} P(S_n \geq y).$$

Por útimo, tomando $x = (y + c)$ obtenemos la desigualdad deseada. ■

Demostración.

Acontinuación se demuestra el Teorema (1.3), y para ello es suficiente mostrar que se cumple la Condición (1.1.9) del Lema (1.1).

Primeramente observemos que

$$P(S_n - S_k \geq -C) \geq q, \qquad para \quad k = 1, \ldots, n-1.$$

Luego, por el Lema (A.1) sabemos que existe un cuantil $k_{1-q}(S_n - S_k)$ tal que

$$k_{1-q}(S_n - S_k) \geq -C, \qquad para \quad k = 1, \ldots, n-1.$$

Entonces tenemos que

$$-k_{1-q}(S_n - S_k) \leq C, \qquad para \quad k = 1, \ldots, n-1.$$

Y como

$$-k_{1-q}(S_n - S_k) = k_q(S_k - S_n),$$

nos queda que

$$k_q(S_k - S_n) \leq C, \qquad para \quad k = 1, \ldots, n-1.$$

Es decir, las hipótesis del Lema (1.1) se satisfacen, y por lo tanto obtenemos la Desigualdad (1.1.8). ■

Para llegar a la desigualdad para la probabilidad del máximo de sumas de variables aleatorias ocupamos el primer teorema de esta sección, el cual no involucra condiciones sobre cuantiles pero en su demostración ocupamos una implicación directa de la definición de cuantil. Asimismo, empleamos un lema el cual si requiere condiciones sobre los cuantiles, y con el cual se puede concluir la demostración de la desigualdad para la probabilidad del máximo de la suma de variables aleatorias independientes.

La cota que obtuvimos para la probabilidad del máximo de sumas es una cota superior que depende de la probabilidad de la suma de las variables aleatorias independietes, esto representa una gran ventaja ya que esta probabilidad se puede aproximar por el Teorema del Límite Central.

1.2 Condiciones sobre Momentos

Teorema 1.4. *Sea $n \geq 2$, $EX_k = 0$, $EX_k^2 < \infty$, para $k = 2, \ldots, n$. Denotemos*

$$b_n = \sum_{k=2}^{n} EX_k^2.$$

Entonces para cada $q \in (0,1)$ y $x \in \mathbb{R}$ tenemos:

$$P\left[\max_{1\leq k\leq n} S_k \geq x\right] \leq \frac{1}{q} P\left[S_n \geq x - \left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}\right] \tag{1.2.1}$$

Demostración.
Observemos que

$$\begin{aligned}
E\left(S_n - S_k\right)^2 &= E\left(\sum_{i=k+1}^{n} X_i\right)^2 \\
&= E\left(\sum_{i=k+1}^{n} X_i^2 + 2\sum_{i\neq j} X_i X_j\right) \\
&= \sum_{i=k+1}^{n} EX_i^2 + 2\sum_{i\neq j} EX_i X_j \\
&= \sum_{i=k+1}^{n} EX_i^2 \\
&\leq b_n.
\end{aligned}$$

Por otro lado, si $1 \leq k \leq n-1$, por la Desigualdad de Chebyshev (A.1) tenemos que

$$P\left[|S_n - S_k| \geq \left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}\right] \leq \frac{1-q}{b_n} E\left(S_n - S_k\right)^2 \leq 1-q,$$

y además, como:

$$P\left[|S_n - S_k| \geq \left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}\right] \geq 1 - P\left[S_n - S_k \geq -\left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}\right],$$

entonces se obtiene que

$$P\left[S_n - S_k \geq -\left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}\right] \geq q, \qquad para \quad k = 1, \ldots, n-1.$$

Por último, aplicando el Teorema (1.3) con $C = \left(\frac{b_n}{1-q}\right)^{\frac{1}{2}}$ se llega a la Desigualdad (1.2.1). ■

Observación 1.1. *Aplicando la Desigualdad* (1.2.1), *con* $q = \frac{1}{2}$, *obtenemos*

$$P\left(\max_{1 \leq x \leq} S_k \geq x\right) \leq 2P\left(S_n \geq x - (2b_n)^{\frac{1}{2}}\right) \tag{1.2.2}$$

para cada $x \in \mathbb{R}$.

Notemos que en este resultado tenemos la ventaja de que la cota superior para la probabilidad del máximo de sumas depende de la probabilidad de la suma de las variables aleatorias.

Sea Y una variable aleatoria discreta definida en $(\Omega, \mathcal{F}, P)$ y sea $B \in \mathcal{F}$ tal que $P(B) > 0$. Se define

$$E\left[Y|B\right] = \sum_y y \frac{P\left(Y = y, B\right)}{P(B)}.$$

Ahora, consideremos a Y una variable aleatoria no negativa, sea $\{B_k\}_{k \in \mathbb{N}}$ una familia de eventos ajenos en $\mathcal{F}$, con $P(B_k) > 0$ para toda $k \in \mathbb{N}$. Entonces, por la definición de esperanza, para $m, n \in \mathbb{N}$, tenemos

$$\begin{aligned}
\sum_{k=m}^{n} P(B_k) E\left[Y | B_k\right] &= \sum_{k=m}^{n} \sum_{y} y P(Y = y, B_k) \\
&= \sum_{y} y \sum_{k=m}^{n} P(Y = y, B_k) \\
&\leq \sum_{y} y P(Y = y) \\
&= E\left[Y\right],
\end{aligned}$$

es decir,

$$E\left[Y\right] \geq \sum_{k=m}^{n} P(B_k) E\left[Y | B_k\right]. \tag{1.2.3}$$

Este resultado es válido también para variables aleatorias en general.

El siguiente teorema trata de ampliar un poco la desigualdad para la probabilidad del máximo de las sumas a la desigualdad para la probabilidad del máximo de los múltiplos de sumas de variables aleatorias independientes.

Teorema 1.5. *Desigualdad de Hàjek-Rényi.*
Sea $EX_k = 0$, $EX_k^2 < \infty$, para $k = 1, \ldots, n$, y sean las constantes $0 < c_n \leq c_{n-1} \leq \cdots \leq c_1$, entonces:

$$P\left[\max_{1 \leq k \leq n} c_k \left|S_k\right| \geq x\right] \leq x^{-2} \left(c_m^2 \sum_{k=1}^{m} EX_k^2 + \sum_{k=m+1}^{n} c_k^2 EX_k^2\right) \tag{1.2.4}$$

para cada $x > 0$ y cada entero positivo $m < n$.

Demostración.
Comenzaremos definiendo la siguiente variable aleatoria

$$Y = \sum_{k=m}^{n-1} \left(c_k^2 - c_{k+1}^2\right) S_k^2 + c_n^2 S_n^2,$$

entonces

$$
\begin{aligned}
EY &= \sum_{k=m}^{n-1} \left(c_k^2 - c_{k+1}^2\right) ES_k^2 + c_n^2 ES_n^2 \\
&= \sum_{k=m}^{n-1} \left(c_k^2 - c_{k+1}^2\right) E\left(\sum_{i=1}^{k} X_i\right)^2 + c_n^2 E\left(\sum_{i=1}^{n} X_i\right)^2 \\
&= \sum_{k=m}^{n-1} \left(c_k^2 - c_{k+1}^2\right) \sum_{i=1}^{k} E\left(X_i^2\right) + c_n^2 \sum_{i=1}^{n} E\left(X_i^2\right) \\
&= c_m^2 \sum_{k=1}^{m} E\left(X_k^2\right) + \sum_{k=m+1}^{n} c_k^2 E\left(X_k^2\right),
\end{aligned}
$$

es decir,

$$
EY = c_m^2 \sum_{k=1}^{m} E\left(X_k^2\right) + \sum_{k=m+1}^{n} c_k^2 E\left(X_k^2\right). \tag{1.2.5}
$$

Ahora definamos los eventos

$$
B_m = \{c_m |S_m| \geq x\},
$$

$$
B_k = \left\{ \max_{m \leq r \leq k-1} c_r |S_r| < x, c_k |S_k| \geq x \right\}, \qquad \textit{para} \quad k = m+1, \ldots, n.
$$

Es claro que

$$
P\left(\max_{m \leq k \leq n} c_k |S_k| \geq x\right) = \sum_{k=m}^{n} P(B_k). \tag{1.2.6}
$$

Por otro lado, aplicando el resultado (1.2.3) en el contexto de este teorema, obtenemos que

$$
EY \geq \sum_{k=m}^{n} P(B_k) E(Y|B_k).
$$

Notemos que si $j > k$, la independencia de las variables aleatorias X_i implican que

$$
E[X_j S_k | B_k] = E[X_j | B_k] E[S_k | B_k],
$$

mientras que

$$
E[X_j | B_k] E[S_k | B_k] = 0,
$$

ya que $EX_j = 0$ para $j = 1, \ldots, n$.

De aquí,

$$\begin{aligned}
E\left(S_j^2 | B_k\right) &\geq E\left(S_k^2 | B_k\right) \\
&= E\left(S_k^2 I_{B_k} | B_k\right) + E\left(S_k^2 I_{B_k^c} | B_k\right) \\
&\geq E\left(S_k^2 I_{B_k} | B_k\right) \\
&= E\left(|S_k|^2 | B_k\right) \\
&\geq \frac{x^2}{c_k^2}.
\end{aligned}$$

Luego, para $k \geq m$ tenemos

$$\begin{aligned}
E\left(Y | B_k\right) &= \sum_{l=m}^{n-1} \left(c_l^2 - c_{l+1}^2\right) E\left(S_l^2 | B_k\right) + c_n^2 E\left(S_n^2 | B_k\right) \\
&\geq \sum_{l=k}^{n-1} \left(c_l^2 - c_{l+1}^2\right) E\left(S_l^2 | B_k\right) + c_n^2 E\left(S_n^2 | B_k\right) \\
&\geq \frac{x^2}{c_k^2} \left\{ \sum_{l=k}^{n-1} \left(c_l^2 - c_{l+1}^2\right) + c_n^2 \right\} \\
&= x^2,
\end{aligned}$$

y por lo tanto, $EY \geq x^2 \sum_{k=m}^{n} P\left(B_k\right)$.

Finalmente, sustituyendo las Igualdades (1.2.5) y (1.2.6) en la ultima desigualdad obtenida, queda desmostrado el teorema. ■

Capítulo 2

Desigualdades con Cotas Exponenciales

A continuación presentamos desigualdades para la probabilidad de la suma de las variables aleatorias independientes. Estas desigualdades representan cotas superiores dadas por una expresión exponencial y se establecen bajo el supuesto de que la función generadora de momentos de cada sumando esta acotada superiormente por una cota exponencial.

Sea el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ donde estan definidas las variables aleatorias independientes $X_1, \ldots, X_n$, y denotemos $S_n = \sum_{k=1}^{n} X_k$.

Teorema 2.1. *Supongamos que existen constantes positivas $g_1, \ldots, g_n$ y T tal que*

$$Ee^{tX_k} \leq e^{g_k \frac{t^2}{2}}, \tag{2.0.1}$$

para $k = 1, \ldots, n$, y $0 \leq t \leq T$, y denotemos $G_n = \sum_{k=1}^{n} g_k$. Entonces

$$P\left(S_n \geq x\right) \leq e^{-\frac{x^2}{2 \cdot G_n}} \qquad si \quad 0 \leq x \leq G_n \cdot T \tag{2.0.2}$$

$$P\left(S_n \geq x\right) \leq e^{-\frac{T \cdot x}{2}} \qquad si \quad x \geq G_n \cdot T \tag{2.0.3}$$

Demostración.

Si $0 < t \leq T$, como e^{tS_n} es una variable aleatoria no negativa, por la Desigualdad de

Markov (A.2) tenemos

$$\begin{aligned} P\left(S_n \geq x\right) &= P\left(e^{tS_n} \geq e^{tx}\right) \\ &= P\left(e^{-tx}e^{tS_n} \geq 1\right) \\ &\leq e^{-tx}Ee^{tS_n}, \end{aligned}$$

para cada $x \in \mathbb{R}$.

Por otro lado, por el hecho de que las variables aleatorias $X_1, \cdots, X_n$ son independientes se obtiene lo siguiente

$$\begin{aligned} Ee^{tS_n} = Ee^{t(X_1+\cdots+X_n)} &= E\left[e^{tX_1}\cdots e^{tX_n}\right] \\ &= Ee^{tX_1}\cdots Ee^{tX_n} = \prod_{k=1}^{n} Ee^{tX_k} \\ &\leq \prod_{k=1}^{n} e^{g_k \frac{t^2}{2}} = e^{(g_1+\cdots+g_n)\frac{t^2}{2}} \\ &= e^{G_n \frac{t^2}{2}}. \end{aligned}$$

De esta manera se sigue

$$P\left(S_n \geq x\right) \leq e^{-tx} \cdot e^{t^2 \frac{G_n}{2}} = e^{t^2 \frac{G_n}{2} - tx}, \tag{2.0.4}$$

para cada x y $0 < t \leq T$.

Ahora, para un valor fijo x, consideremos la función

$$f\left(t\right) = \frac{G_n}{2}t^2 - xt,$$

y minimizaremos la función $e^{f(t)}$, lo que corresponde a minimizar la función $f\left(t\right)$, ya que la función exponencial es creciente.

Si $0 < x \leq G_nT$, entonces la ecuación $f'\left(t\right) = 0$ tiene una única solución en $t = \frac{x}{G_n}$, que corresponde al valor donde alcanza su mínimo la función $f\left(t\right)$. Notemos que esta solución satisface la condición $0 < t \leq T$, por lo cual podemos sustituir el valor de $t = \frac{x}{G_n}$ en la Desigualdad (2.0.4) y de aquí obtenemos la Desigualdad (2.0.2).

En este caso podemos observar que si $x = 0$ también se cumple la desigualdad (2.0.2).

Si $x \geq G_nT$, entonces $f'\left(t\right) = G_nt - x \leq 0$, y así la fución $f\left(t\right)$ es decreciente. Por consiguiente, la función $e^{f(t)}$ alcanza su mínimo cuando $t = T$. Sustiyendo el valor donde se alcanza el mínimo en la Desigualdad (2.0.4) obtenemos la Desigualdad (2.0.3).

■

Notemos que el teorema establece las desigualdades para las probabilidades de la suma de variables sólo para valores no negativos de la suma de las variables. La siguiente observación nos sirve para ampliar estas desigualdes a los valores negativos de la suma de las variables.

Observación 2.1. *Notemos que si la Condición* (2.0.1) *se cumple para* $-T \leq t \leq 0$ *y para algunas constantes positivas* $T, g_1, \ldots, g_n$*, entonces podemos obtener de manera análoga a la demostración del Teorema* (2.1) *las siguientes desigualdades:*

$$P\left(S_n \leq -x\right) \leq e^{\frac{-x^2}{2G_n}} \qquad si \quad 0 \leq x \leq G_n T. \tag{2.0.5}$$

$$P\left(S_n \leq -x\right) \leq e^{\frac{-Tx}{2}} \qquad si \quad x \geq G_n T. \tag{2.0.6}$$

Para demostración de las desigualdades anteriores cabe notar que si las variables aleatorias $X_1, \ldots, X_n$ *satisfacen la Condición* (2.0.1) *del Teorema* (2.1) *para* $-T \leq t \leq 0$*, entonces las variables aleatorias* $-X_1, \ldots, -X_n$ *también la satisfacen pero para* $0 \leq t \leq T$*, y por tanto se cumplen las Desigualdades* (2.0.2) *y* (2.0.3) *para la variable aleatoria* $-S_n$.

Los resultados del Teorema (2.1) junto con la Observación (2.1) nos llevan al siguiente teorema.

Teorema 2.2. *Supongamos que existen constantes positivas* $g_1, \ldots, g_n$ *y* T *tal que*

$$Ee^{tX_k} \leq e^{g_k \frac{t^2}{2}}$$

para $k = 1, \ldots, n$, *y* $|t| \leq T$.

Entonces se cumplen las Desigualdades (2.0.2), (2.0.3), (2.0.5) *y* (2.0.6).

Veamos un ejemplo en el que los sumandos corresponden a variables aleatorias normales, y aunque de antemano sabemos que la suma de variables aleatorias independientes con distribución normal también tiene una distribución normal, el teorema anterior nos ayuda a estimar la probabilidad sin necesidad de calcular su distribución y la integral para obtener la probabilidad deseada.

Ejemplo 2.1. *Supongamos que* $X_k \sim N(0, \sigma_k)$, *para* $k = 1, \ldots, n$. *Entonces*

$$Ee^{tX_k} = e^{\sigma_k^2 \frac{t^2}{2}},$$

y así la Condición (2.0.1) *se cumple para* $g_k = \sigma_k^2$ *y para cada* $T > 0$.
En este caso tenemos que $G_n = \sum_{k=1}^n \sigma_k^2$, *y por el teorema* (2.1) *obtenemos* $P(S_n \geq x) \leq e^{-\frac{x^2}{2G_n}}$, *para cada* $x \geq 0$.

Teorema 2.3. *Desigualdes de Bernstein.*
Supongamos que $EX_k = 0, \sigma_k^2 = EX_k^2 < \infty$, *para* $k = 1, \ldots, n$, *y* $B_n = \sum_{k=1}^n \sigma_k^2$. *Además supongamos que existe una constante positiva* H *tal que:*

$$|EX_k^m| \leq \frac{1}{2} m! \sigma_k^2 H^{m-2}, \qquad (k = 1, \ldots, n.), \tag{2.0.7}$$

para todos lo enteros $m \geq 2$.
Entonces

$$P(S_n \geq x) \leq e^{\frac{-x^2}{4B_n}} \quad si \quad 0 \leq x \leq \frac{B_n}{H} \tag{2.0.8}$$

$$P(S_n \geq x) \leq e^{\frac{-x}{4H}} \quad si \quad x \geq \frac{B_n}{H} \tag{2.0.9}$$

$$P(S_n \leq -x) \leq e^{\frac{-x^2}{4B_n}} \quad si \quad 0 \leq x \leq \frac{B_n}{H} \tag{2.0.10}$$

$$P(S_n \leq -x) \leq e^{\frac{-x}{4H}} \quad si \quad x \geq \frac{B_n}{H} \tag{2.0.11}$$

Demostración.
Consideremos la identidad

$$Ee^{tX_k} = 1 + t \cdot EX_k + \frac{t^2}{2}\sigma_k^2 + \frac{t^3}{6}EX_k^3 + \cdots$$

para $k = 1, \ldots, n$. Luego, como $EX_k = 0$, tenemos que

$$Ee^{tX_k} = 1 + \frac{t^2}{2}\sigma_k^2 + \frac{t^3}{6}EX_k^3 + \cdots \leq 1 + \frac{1}{2}t^2\sigma_k^2\left(1 + H|t| + H^2t^2 + \cdots\right), \tag{2.0.12}$$

donde la desigualdad se sigue del hecho de que

$$\begin{aligned}
\frac{t^m}{m!}EX_k^m &\leq \frac{|t^m|}{m!}|EX_k^m| \\
&\leq \frac{|t^m|}{m!}\left(\frac{m!}{2}\sigma_k^2 H^{m-2}\right) \\
&= \frac{|t^m|}{2}\sigma_k^2 H^{m-2},
\end{aligned}$$

lo cual implica

$$\frac{t^m}{m!}EX_k^m \leq \frac{|t^m|}{2}\sigma_k^2 H^{m-2}.$$

Ahora, si $|t| \leq \frac{1}{2H}$ entonces $H|t| \leq \frac{1}{2}$, y por lo tanto

$$1 + H|t| + H^2t^2 + \cdots = \frac{1}{1-|t|H}.$$

Sustituyendo la igualdad anterior en la Desigualdad (2.0.12) nos queda

$$1 + \frac{t^2}{2}\sigma_k^2\left(\frac{1}{1-|t|H}\right) \leq 1 + \sigma_k^2 t^2 \leq e^{\sigma_k^2 t^2}.$$

Resumiendo tenemos hasta aquí que

$$Ee^{tX_k} \leq e^{\sigma_k^2 t^2}$$

para $|t| \leq \frac{1}{2H}$.

Luego, denotando $g_k = 2\sigma_k^2$ tenemos que $G_n = \sum_{k=1}^n 2\sigma_k^2 = 2B_n$, lo cual implica que se cumplen las hipótesis del Teorema (2.2), lo que nos conduce fácilmente a obtener las desigualdades deseadas. ■

La condición del teorema claramente no es sobre la función generadora de momentos de cada sumando, como hemos visto en resultados anteriores, sin embargo lo implica y con ello se pudo concluir la demostración del teorema.

Observación 2.2. *La Condición* (2.0.7) *es conocida como la condición de Bernstein. Acontinuación presentamos dos condiciones suficientes para la condición de Bernstein.*

Condición 1. *Existe una constante positiva C tal que $P(|X_k| \leq C) = 1$ y $E(X_k) = 0$, para $k = 1, \ldots, n$. En este caso observemos que*

$$|EX_k^m| = \left|\int_{-C}^{C} x^m dV_k(x)\right| \leq C^{m-2}\int_{-C}^{C} x^2 dV_k(x) = C^{m-2}\sigma_k^2$$

para toda $m \geq 2$, donde $V_k(x)$ es la función de distribucón de la variable aleatoria X_k. Entonces la Condición (2.0.7) *se cumple con $H = C$.*

Condición 2. *Las variables aleatorias $X_1, \ldots, X_n$ son independientes e idénticamente distribuidas que satisfacen la condición de Cramér. Es decir, existe la función generadora de momentos de cada variable aleatoria X_k en algún intervalo con centro en el origen. En este caso, esta condición implica que existen los momentos de todos los ordenes de cada variable aleatoria X_k, y por lo tanto se cumple la condición de Bernstein.*

Capítulo 3

Desigualdades para Momentos de Sumas

Los momentos de una variable aleatoria X nos permiten tener una mejor idea sobre las características de la distribución de la variable. El ejemplo más aplicado es la varianza de una variable aleatoria, la cual se expresa como el segundo momento con respecto a la media y nos describe, en promedio, la acumulación o dispersión de los valores de la variable aleatoria con respecto a su media. Otro ejemplo es la medida de asimetría de X, dada por

$$E\left[X - E(X)\right]^3,$$

la cual nos sirve para estimar la falta de simetría de la distribución de la variable.

En este capítulo presentamos desigualdades para los momentos de una suma de variables aleatorias independientes, en particular para momentos mayores que el primero, ya que en todos los resultados que se exponen la hipótesis es que cada sumando tiene media cero. De esta manera, el segundo momento de la variable aleatoria representa la varianza de la misma y, asimismo, el tercer momento representa la medida de asimetría de la variable aleatoria, etc.

Consideremos un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ donde estan definidas las variables aleatorias $X_1, \ldots, X_n$, no necesariamente independientes, entonces:

$$E\left|S_n\right|^p \leq \sum_{k=1}^{n} E|X_k|^p \qquad si \quad 0 < p < 1, \tag{3.0.1}$$

$$E|S_n|^p \leq n^{p-1} \sum_{k=1}^{n} E|X_k|^p \quad si \quad p > 1, \tag{3.0.2}$$

donde $S_n = \sum_{k=1}^{n} X_k$.

Nótese que aunque algún momento $E|X_k|^p$ sea infinito las desigualdades anteriores siguen siendo válidas. En el caso en que ningún momento sea infinito las Desigualdades (3.0.1) y (3.0.2) se derivan de las siguientes desigualdades:

$$\left|\sum_{k=1}^{n} a_n\right|^p \leq \sum_{k=1}^{n} |a_n|^p \quad si \quad 0 < p \leq 1,$$

$$\left|\sum_{k=1}^{n} a_n\right|^p \leq n^{p-1} \sum_{k=1}^{n} |a_n|^p \quad si \quad p > 1,$$

que se cumplen para cada entero positivo n y números reales arbitrarios $a_1, \ldots, a_n$.

En adelante supondrémos que $X_1, \ldots, X_n$ son variables aleatorias independientes.

Teorema 3.1. *(Desigualdad de Rusenthal)*
Sean $X_1, \ldots, X_n$ variables aleatorias independientes y Supongamos que $EX_k = 0$, para $k = 1, \ldots, n$, y sea $p \geq 2$. Denotemos

$$M_{p,n} = \sum_{k=1}^{n} E|X_k|^p, \qquad B_n = \sum_{k=1}^{n} EX_k^2.$$

Entonces

$$E|S_n|^p \leq C(p)\left(M_{p,n} + B_n^{\frac{p}{2}}\right) \tag{3.0.3}$$

donde $C(p)$ es una constante positiva que depende sólo de p.

La demostración del Teorema (3.1) es consecuencia de los siguientes resultados.

Lema 3.1. *Sean $X_1, \ldots, X_n$ variables aleatorias independientes con funciones de distribución $V_1(x), \ldots, V_n(x)$ respectivamente, y sea $\boldsymbol{y}$ numero positivo. Denotemos*

$$\mu(-\infty, \boldsymbol{y}) = \sum_{k=1}^{n} \int_{x<y} x dV_k(x)$$

$$B(-\infty, y) = \sum_{k=1}^{n} \int_{x<y} x^2 dV_k(x).$$

Entonces

$$P(S_n \geq x) \leq \sum_{k=1}^{n} P(X_k \geq y) + exp\left\{\frac{x}{y} - \left(\frac{x - \mu(-\infty, y)}{y} + \frac{B(-\infty, y)}{y^2}\right) log\left(1 + \frac{xy}{B(-\infty, y)}\right)\right\}$$

para cada $x > 0$.

Demostración.

Consideremos las siguientes variables aleatorias:

$$Z_k = \begin{cases} X_k & si \quad X_k < \mathbf{y} \\ 0 & si \quad X_k \geq \mathbf{y} \end{cases}$$

y

$$T_n = \sum_{k=1}^{n} Z_k.$$

Si $S_n \geq x$ entonces $\sum_{k=1}^{n} X_k \geq x$, para cada $x \in \mathbb{R}$. Además si $X_k < \mathbf{y}$ para toda $k = 1, \ldots, n$, entonces $\sum_{k=1}^{n} X_k = \sum_{k=1}^{n} Z_k = T_n \geq x$. En cambio, si $X_k \geq \mathbf{y}$ para toda $k = 1, \ldots, n$, entonces $\sum_{k=1}^{n} Z_k = T_n = 0$, lo cual es diferente al valor de S_n.

De aquí concluimos que

$$\{S_n \geq x\} \subset \{T_n \geq x\} \cup \{S_n \neq T_n\},$$

es decir

$$P(S_n \geq x) \leq P(T_n \geq x) + P(S_n \neq T_n) \tag{3.0.4}$$

para cada $x \in \mathbb{R}$.

Por otro lado:

$$\begin{aligned} P(T_n \geq x) &= P\left(e^{T_n} \geq e^x\right) \\ &= P\left(e^{hT_n} \geq e^{hx}\right), \quad para \quad h > 0, \\ &= P\left(e^{h(T_n - x)} \geq 1\right) \\ &\leq Ee^{h(T_n - x)} \\ &= e^{-hx} Ee^{hT_n}, \end{aligned} \tag{3.0.5}$$

y esta manera, juntando la desigualdad (3.0.4) y (3.0.5) nos queda la siguiente desigualdad

$$P(S_n \geq x) \leq e^{-hx} Ee^{hT_n} + \sum_{k=1}^{n} P(X_k \geq \mathbf{y}). \tag{3.0.6}$$

Luego, es fácil mostrar que la función

$$\frac{e^x - 1 - x}{x^2}$$

es no decreciente para $x > 0$, lo cual implica que para $x > 0$ la función

$$\frac{e^x - 1 - hx}{x^2}, \quad \textit{donde} \quad h > 0,$$

es también no decreciente para $x > 0$.

Además, notemos lo siguiente:

$$\begin{aligned}
Ee^{hZ_k} &= \int_{-\infty}^{\mathbf{y}} e^{hx} dV_k(x) + \int_{\mathbf{y}}^{\infty} e^{h \cdot 0} dV_k(x) \\
&= \int_{-\infty}^{\mathbf{y}} e^{hx} dV_k(x) + \int_{\mathbf{y}}^{\infty} dV_k(x) \\
&= \int_{-\infty}^{\mathbf{y}} e^{hx} dV_k(x) + 1 - \int_{-\infty}^{\mathbf{y}} dV_k(x) \\
&= 1 + \int_{-\infty}^{\mathbf{y}} \left(e^{hx} - 1\right) dV_k(x) \qquad (3.0.7) \\
&= 1 + \int_{-\infty}^{\mathbf{y}} \left(e^{hx} - 1 + hx - hx\right) dV_k(x) \\
&= 1 + \int_{-\infty}^{\mathbf{y}} \left(e^{hx} - 1 - hx\right) dV_k(x) + h \int_{-\infty}^{\mathbf{y}} x dV_k(x) \\
&\leq 1 + h \int_{-\infty}^{\mathbf{y}} x dV_k(x) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right) \mathbf{y}^{-2} \int_{-\infty}^{\mathbf{y}} x^2 dV_k(x).
\end{aligned}$$

La última desigualdad se debe a que tenemos que $x \leq \mathbf{y}$, entonces

$$\frac{e^{hx} - 1 + hx}{x^2} \leq \frac{e^{h\mathbf{y}} - 1 + h\mathbf{y}}{\mathbf{y}^2},$$

y así

$$x^2 \frac{e^{hx} - 1 + hx}{x^2} \leq x^2 \frac{e^{h\mathbf{y}} - 1 + h\mathbf{y}}{\mathbf{y}^2}.$$

También observemos que las variables aleatorias $Z_1, \ldots, Z_n$ son independientes ya que están en función de las variables aleatorias $X_1, \ldots, X_n$, las cuales por hipótesis son independientes. De esta manera:

$$\begin{aligned}
e^{-hx} E e^{nT_n} &= e^{-hx} E \left(e^{hZ_1} \cdots e^{hZ_n}\right) \\
&= e^{-hx} \prod_{k=1}^{n} E e^{hZ_k}.
\end{aligned}$$

Por la cota (3.0.7), obtenida anteriormente para Ee^{hZ_k}, se sigue que:

$$\prod_{k=1}^{n} Ee^{hZ_k} \leq \prod_{k=1}^{n} \left(1 + h\int_{-\infty}^{\mathbf{y}} x dV_k(x) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}\int_{-\infty}^{\mathbf{y}} x^2 dV_k(x)\right).$$

Denotando $w_k = h\int_{-\infty}^{\mathbf{y}} x dV_k(x) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}\int_{-\infty}^{\mathbf{y}} x^2 dV_k(x)$, podemos abreviar la desigualdad anterior como sigue:

$$\begin{aligned}
\prod_{k=1}^{n} Ee^{hZ_k} &\leq \prod_{k=1}^{n} (1 + w_k) \\
&\leq \prod_{k=1}^{n} e^{w_k} \\
&= e^{\sum_{k=1}^{n} w_k} \\
&= e^{h\sum_{k=1}^{n}\int_{x\leq \mathbf{y}} x dV_k(x) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}\sum_{k=1}^{n}\int_{x\leq \mathbf{y}} x^2 dV_k(x)} \\
&= e^{h\mu(-\infty,\mathbf{y}) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}B(-\infty,\mathbf{y})}.
\end{aligned}$$

Vamos a multiplicar la expresión anterior por e^{-xh} en ambos lados de la desigualdad, lo cual queda como sigue:

$$e^{-xh}\prod_{k=1}^{n} Ee^{hZ_k} \leq e^{-xh} e^{h\mu(-\infty,\mathbf{y}) + \left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}B(-\infty,\mathbf{y})},$$

para luego tomar el valor de h de la siguiente manera $h = \frac{1}{\mathbf{y}} log\left(1 + \frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right)$, ya que así tenemos la siguiente equivalencia:

$$e^{h\mathbf{y}} - 1 = \frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}. \tag{3.0.8}$$

Luego, mediante la resta de $h\mathbf{y}$ y la multiplicación de $\mathbf{y}^{-2}B(-\infty,\mathbf{y})$ en ambos lados de la igualdad de (3.0.8) llegamos a la siguiente igualdad:

$$\begin{aligned}
\left(e^{h\mathbf{y}} - 1 - h\mathbf{y}\right)\mathbf{y}^{-2}B(-\infty,\mathbf{y}) &= \left(\frac{x\mathbf{y}}{B(-\infty,\mathbf{y})} - h\mathbf{y}\right)\mathbf{y}^{-2}B(-\infty,\mathbf{y}) \\
&= \frac{x}{\mathbf{y}} - \frac{B(-\infty,\mathbf{y})}{\mathbf{y}^2} log\left(1 + \frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right).
\end{aligned}$$

Hasta aqui hemos mostrado

$$\begin{aligned}
e^{-hx}Ee^{hT_n} &= e^{\left\{-hx\prod_{k=1}^{n} Ee^{hZ_k}\right\}} \\
&\leq e^{\left\{-\frac{x}{\mathbf{y}} log\left(1+\frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right) + \frac{\mu(-\infty,\mathbf{y})}{\mathbf{y}} log\left(1+\frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right) + \frac{x}{\mathbf{y}} - \frac{B(-\infty,\mathbf{y})}{y^2} log\left(1+\frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right)\right\}} \\
&= e^{\left\{\frac{x}{\mathbf{y}} - \left(\frac{B(-\infty,\mathbf{y})}{\mathbf{y}^2} - \frac{\mu(-\infty,\mathbf{y})}{\mathbf{y}} + \frac{x}{\mathbf{y}}\right) log\left(1+\frac{x\mathbf{y}}{B(-\infty,\mathbf{y})}\right)\right\}}
\end{aligned}$$

(3.0.9)

Por último, combinando la Desigualdad (3.0.9) con la Desigualdad (3.0.6) se llega a la desigualdad deseada. ∎

Observación 3.1.

- *Por hipótesis tenemos que $EX_k = 0$ para todos los valores de k, por lo cual $\mu(-\infty, \mathbf{y}) \leq 0$, es decir $-\mu(-\infty, \mathbf{y}) \geq 0$.*

- *Por la definición de B_n se sigue que $0 < B(-\infty, \mathbf{y}) \leq B_n$.*

  Con el punto anterior y este llegamos a la siguiente desigualdad:

  $$\frac{x}{\mathbf{y}} - \left(\frac{B(-\infty, \mathbf{y})}{\mathbf{y}^2} - \frac{\mu(-\infty, \mathbf{y})}{\mathbf{y}} + \frac{x}{\mathbf{y}}\right) log\left(1 + \frac{x\mathbf{y}}{B(-\infty, \mathbf{y})}\right) \leq \frac{x}{\mathbf{y}} - \left(\frac{x}{\mathbf{y}}\right) log\left(1 + \frac{x\mathbf{y}}{B_n}\right).$$

- *Por el resultado del Lema (3.1) y por el punto anterior se obtiene lo siguiente:*

  $$\begin{aligned} P(S_n \geq x) &\leq \sum_{k=1}^{n} P(X_k \geq \mathbf{y}) + exp\left\{\frac{x}{\mathbf{y}} - \left(\frac{x - \mu(-\infty, \mathbf{y})}{\mathbf{y}} + \frac{B(-\infty, \mathbf{y})}{\mathbf{y}^2}\right) log\left(1 + \frac{x\mathbf{y}}{B(-\infty, \mathbf{y})}\right)\right\} \\ &\leq \sum_{k=1}^{n} P(X_k \geq \mathbf{y}) + exp\left\{\frac{x}{\mathbf{y}} - \left(\frac{x}{\mathbf{y}}\right) log\left(1 + \frac{x\mathbf{y}}{B_n}\right)\right\}, \end{aligned}$$

  y de aquí

  $$P(|S_n| \geq x) \leq \sum_{k=1}^{n} P(|X_k| \geq \mathbf{y}) + 2exp\left\{\frac{x}{\mathbf{y}} - \left(\frac{x}{\mathbf{y}}\right) log\left(1 + \frac{x\mathbf{y}}{B_n}\right)\right\} \qquad (3.0.10)$$

  para cada x y $\mathbf{y}$ positivos.

Como ejemplo del Lema (A.2), tomemos $g(x) = |X|^p$, donde $p > 0$. Esta función claramente satisface las condiciones del Lema (A.2), por lo cual tenemos que

$$E|X|^p = p\int_0^{\infty} P(|X| \geq x)x^{p-1}dx, \qquad (3.0.11)$$

para una variable aleatoria albitraria X.

Acontinuación se presenta la demostración del Teorema (3.1).

Demostración.

Sustituyamos $\mathbf{y} = \frac{x}{r}$, con $r > \frac{p}{2}$, en la Desigualdad (3.0.10), con lo cual resulta:

$$P(|S_n| \geq x) \leq \sum_{k=1}^{n} P(|X_k| \geq \frac{x}{r}) + 2 \cdot exp\left\{r - rlog\left(1 + \frac{x^2}{rB_n}\right)\right\}.$$

Ahora multiplicando por px^{p-1} en ambos lados de la desigualdad anterior y después integrando sobre la parte positiva de la recta real se obtiene lo siguiente:

$$\int_0^\infty P(|S_n| \geq x)px^{p-1}dx \leq p\int_o^\infty \sum_{k=1}^n P(|X_k| \geq \frac{x}{r})x^{p-1}dx + 2p\int_0^\infty exp\left\{r - rlog\left(1+\frac{x^2}{rB_n}\right)\right\}x^{p-1}dx.$$

Aplicando la Igualdad (3.0.11), obtenemos

$$\begin{aligned} E|S_n|^p &\leq p\int_o^\infty \sum_{k=1}^n P(|X_k| \geq \frac{x}{r})x^{p-1}dx + 2p\int_0^\infty exp\left\{r - rlog\left(1+\frac{x^2}{rB_n}\right)\right\}x^{p-1}dx \\ &= p\int_o^\infty \sum_{k=1}^n P(|rX_k| \geq x)x^{p-1}dx + 2p\int_0^\infty e^r e^{-rlog\left(1+\frac{x^2}{rB_n}\right)}x^{p-1}dx \\ &= \sum_{k=1}^n E|rX_k|^p + 2pe^r\int_0^\infty \frac{x^{p-1}}{\left(1+\frac{x^2}{rB_n}\right)^r}dx \\ &= r^p\sum_{k=1}^n E|X_k|^p + 2pe^r\int_0^\infty \frac{x^{p-1}}{\left(1+\frac{x^2}{rB_n}\right)^r}dx. \end{aligned}$$

Para simplificar esta expresión podemos usar la identidad obtenida de [4]:

$$\int_0^\infty \frac{x^{p-1}}{\left(1+\frac{x^2}{rB_n}\right)^r}dx = \frac{r^{\frac{p}{2}}B_n^{\frac{p}{2}}}{2}B\left(\frac{p}{2}, r-\frac{p}{2}\right),$$

donde $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$ es la función Beta.

Por lo tanto, sustituyendo nos queda que:

$$E|S_n|^p \leq r^p\sum_{k=1}^n E|X_k|^p + pr^{\frac{p}{2}}B_n^{\frac{p}{2}}B\left(\frac{p}{2}, r-\frac{p}{2}\right), \tag{3.0.12}$$

y tomando $C(p) = pr^pe^rB\left(\frac{p}{2}, r-\frac{p}{2}\right)$, se obtiene la desigualdad deseada del Teorema (3.1).

∎

Teorema 3.2. *Sean $X_1, \ldots, X_n$, variables aleatorias independientes con media cero, y sea $p \geq 2$. Entonces*

$$E|S_n|^p \leq C(p)n^{\frac{p}{2}-1}\sum_{k=1}^n E|X_k|^p, \tag{3.0.13}$$

donde $C(p)$ es una constante positiva que depende sólo de p.

Demostración.

Primeramente apliquemos la Desigualdad de Lyapunov (A.3) a la variable aleatoria S_n, como sigue:

$$(E|S_n|^2)^{\frac{1}{2}} \leq (E|S_n|^p)^{\frac{1}{p}},$$

para cada $p \geq 2$. Luego, por un lado tenemos lo siguiente:

$$\begin{aligned}(E|S_n|^2)^{\frac{1}{2}} &= (E|X_1 + \cdots + X_n|^2)^{\frac{1}{2}} \\ &= \left(\sum_{k=1}^{n} E|X_k|^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{k=1}^{n} EX_k^2\right)^{\frac{1}{2}} \\ &= (B_n)^{\frac{1}{2}},\end{aligned}$$

y además

$$\begin{aligned}(E|S_n|^p)^{\frac{1}{p}} &= (E|X_1 + \cdots + X_n|^p)^{\frac{1}{p}} \\ &= \left(\sum_{k=1}^{n} E|X_k|^p\right)^{\frac{1}{p}} \\ &= (M_{n,p})^{\frac{1}{p}},\end{aligned}$$

donde B_n y $M_{n,p}$ están definidos como en el Teorema (3.1), y de este modo tenemos que

$$(B_n)^{\frac{1}{2}} \leq (M_{p,n})^{\frac{1}{p}},$$

es decir,

$$\left(\frac{B_n}{n}\right)^{\frac{p}{2}} \leq \left(\frac{M_{p,n}}{n}\right),$$

y en consecuencia

$$B_n^{\frac{p}{2}} \leq n^{\frac{p}{2}-1} M_{p,n}.$$

Por último, tomando en cuenta el resultado del Teorema (3.1) llegamos a que

$$\begin{aligned}E|S_n|^p &\leq C(p)\left(M_{p,n} + n^{\frac{p}{2}-1} M_{p,n}\right) \\ &\leq C(p)\left(n^{\frac{p}{2}-1} M_{p,n} + n^{\frac{p}{2}-1} M_{p,n}\right) \\ &= 2n^{\frac{p}{2}-1} C(p) M_{p,n},\end{aligned}$$

con lo cual se concluye la demostración. ■

Observación 3.2.

- *La Desigualdad (3.0.13) es muy parecida a la Desigualdad (3.0.2), sin embargo la Desigualdad (3.0.2) es establecida sin la condición de independencia.*

- *Podemos obtener una mejor desigualdad si en lugar de basarnos en la Desigualdad de Lyapunov (A.3) en el Teorema (3.2) usamos el Teorema (A.1).*

  En efecto, sea Y una variable aleatoria con función de distribución

$$F_Y(x) = P(Y < x) = \frac{1}{n}\sum_{k=1}^{n} P(X_k < x).$$

  De aquí

$$E|Y|^r = \frac{1}{n}\sum_{k=1}^{n} E|X_k|^r,$$

  para cualquier $r > 0$, y además

$$P(Y \neq 0) = \frac{1}{n}\sum_{k=1}^{n} P(X_k \neq 0).$$

  Ahora, aplicando el Teorema (A.1) a la variable aleatoria Y, con $r = 2$ y $s = p \geq 2$, obtenemos

$$\begin{aligned}(E|Y|^2)^{\frac{1}{2}} &= \left(\frac{1}{n}\sum_{k=1}^{n} E|X_k|^2\right)^{\frac{1}{2}} \\ &\leq \left(\sum_{k=1}^{n} P(X_k \neq 0)\right)^{\frac{1}{2}-\frac{1}{p}} (E|Y|^p)^{\frac{1}{p}} \\ &= \left(\sum_{k=1}^{n} P(X_k \neq 0)\right)^{\frac{1}{2}-\frac{1}{p}} \left(\frac{1}{n}\sum_{k=1}^{n} E|X_k|^p\right)^{\frac{1}{p}}.\end{aligned}$$

  Notemos que

$$\left(\frac{1}{n}\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} E|X_k|^2\right)^{\frac{1}{2}} \leq \left(\frac{1}{n}\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} P(X_k \neq 0)\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{k=1}^{n} E|X_k|^p\right)^{\frac{1}{p}}$$

  y

$$\begin{aligned}\left(\sum_{k=1}^{n} E|X_k|^2\right)^{\frac{1}{2}} &\leq \frac{(\frac{1}{n})^{\frac{1}{p}}}{(\frac{1}{n})^{\frac{1}{2}}} \left(\sum_{k=1}^{n} P(X_k \neq 0)\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{k=1}^{n} E|X_k|^p\right)^{\frac{1}{p}} \\ &\leq \left(\sum_{k=1}^{n} P(X_k \neq 0)\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{k=1}^{n} E|X_k|^p\right)^{\frac{1}{p}}.\end{aligned}$$

Por lo cual

$$B_n^{\frac{1}{2}} \leq \left(\sum_{k=1}^{n} P(X_k \neq 0) \right)^{\frac{1}{2}-\frac{1}{p}} M_{p,n}^{\frac{1}{p}}. \tag{3.0.14}$$

El siguiente resultado es una consecuencia de la Observación (3.2) y de la Desigualdad (3.0.3) del Teorema (3.1).

Teorema 3.3. *Sean $X_1, \ldots, X_n$, variables aleatorias independientes, con $EX_k = 0$ para cada $k = 1, \ldots, n$, y sea $p \geq 2$. Entonces*

$$E|S_n|^p \leq C(p) \left[1 + \left(\sum_{k=1}^{n} P(X_k \neq 0) \right)^{\frac{p}{2}-1} \right] M_{p,n} \tag{3.0.15}$$

Demostración.
Elevando a la p en ambos lados de la Desigualdad (3.0.14) resulta que

$$B_n^{\frac{p}{2}} \leq \left(\sum_{k=1}^{n} P(X_k \neq 0) \right)^{\frac{p}{2}-1} M_{p,n}.$$

Por otro lado, aplicando el Teorema (3.1), la Desigualdad (3.0.3) se cumple, y por lo tanto

$$\begin{aligned} E|S_n|^p &\leq C(p) \left[M_{p,n} + B_n^{\frac{p}{2}} \right] \\ &\leq C(p) \left[M_{p,n} + \left(\sum_{k=1}^{n} P(X_k \neq 0) \right)^{\frac{p}{2}-1} M_{p,n} \right] \\ &= C(p) M_{p,n} \left[1 + \left(\sum_{k=1}^{n} P(X_k \neq 0) \right)^{\frac{p}{2}-1} \right]. \end{aligned}$$

■

Observación 3.3.
Si la suma $\sum_{k=1}^{n} P(X_k \neq 0)$ crece más lento que n, entonces la Desigualdad (3.0.15) del Teorema (3.3) nos proporciona una mejor estimación que la Desigualdad (3.0.13) del Teorema (3.2).

Parte II

Desigualdades tipo Bonferroni y Desigualdades para Probabilidades de Colas

Capítulo 4

Desigualdades tipo Bonferroni

4.1 Introducción

Sea el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ donde están definidos los eventos y las variables aleatorias que estaremos usando a lo largo de este capítulo.

Sea $\mathcal{A} = \{A_1, \ldots, A_n\}$ una colección de eventos y definamos la variable aleatoria ν_n que representa el número de eventos A_j's, $1 \leq j \leq n$, que ocurren. Además, se define $S_0 = S_{0,n} = 1$, y para $k \geq 1$,

$$S_k = S_{k,n} = \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} P\left(A_{i_1}, \cdots, A_{i_k}\right) \tag{4.1.1}$$

donde $A_j, A_\ell = A_j \cap A_\ell$.

Por otro lado, para $m \geq 0$, denotemos

$$P_{(m)} = P(\nu_n \geq m) \quad y \quad P_{[m]} = P(\nu_n = m). \tag{4.1.2}$$

En términos generales, a las desigualdades de la forma

$$\sum_{t \geq 0} a_t(n) S_t \leq P_{(m)} \leq \sum_{t \geq 0} b_t(n) S_t$$

y

$$\sum_{t \geq 0} a_t(n) S_t \leq P_{[m]} \leq \sum_{t \geq 0} b_t(n) S_t$$

para ciertos coeficientes $a_t(n)$ y $b_t(n)$, se les conoce como **Desigualdades tipo Bonferroni.**

Este tipo de desigualdades, en su forma más simple, aparecen en la teoría básica de probabilidad. En efecto, observemos que si $n = 1$, es decir cuando $\mathcal{A} = \{A_1\}$, entonces

$$P_{[0]} = P(\nu_n = 0) = P(\nu_1 = 0) = P(A_1^c) = 1 - P(A_1) = S_0 - S_1,$$

y para $n = 2$, tenemos

$$\begin{aligned} P_{(1)} &= P(\nu_2 \geq 1) \\ &= P(A_1) + P(A_2) - P(A_1 A_2) \\ &= S_1 - S_2, \end{aligned}$$

lo cual implica que

$$P_{(1)} \leq S_1.$$

Luego, por las propiedades de la medida de probabilidad tenemos que para $n \geq 2$

$$P\left(A_1 \cup A_2\right) \leq P(A_1) + P(A_2) \leq \sum_{j=1}^{n} P(A_j) = S_{1,n}.$$

Aplicando argumentos de inducción podemos mostrar que se cumple

$$P_{(1)} = P\left(\nu_n \geq 1\right) \leq S_{1,n}, \quad n \in \mathbb{N}.$$

El objetivo del presente capítulo es desarrollar desigualdades tipo Bonferroni más generales para valores arbitrarios de n y m. Para esto aplicaremos un método básico para obtenerlas, al cual se le conoce como el Método de las Indicadoras. Éste método surge del Álgebra Booleana y, en general, consiste en aplicar técnicas de análisis combinatorio para establecer desigualdades no probabilísticas que sean equivalentes a desigualdades de nuestro interés.

Al final del capítulo expondremos algunas ideas generales sobre las posibles aplicaciones de este tipo de desigualdades.

4.2 El Método de las indicadoras

Sea $I(C)$ la variable indicadora del evento C, esto es:

$$I(C) = \begin{cases} 1 & si \quad C \quad ocurre. \\ 0 & si \quad C \quad no \quad ocurre. \end{cases} \tag{4.2.1}$$

Observemos que se satisface lo siguiente:

$$I(A_{i_1}, A_{i_2}, \cdots A_{i_k}) = I(A_{i_1})I(A_{i_2}) \cdots I(A_{i_k}), \tag{4.2.2}$$

$$I(C^c) = 1 - I(C), \tag{4.2.3}$$

y

$$E\left[I(C)\right] = P(C), \tag{4.2.4}$$

donde E es el operador esperanza.
Por otro lado,

$$S_k = E\left[J_k\right], \tag{4.2.5}$$

donde $J_k = \sum I(A_{i_1})I(A_{i_2}) \cdots I(A_{i_k})$, con la sumatoria sobre $1 \leq i_1 \leq \cdots \leq i_k \leq n$.

Aplicando las propiedades de linealidad y monotonía de la esperanza, tenemos que la desigualdad

$$\sum_{k=0}^{n} c_k S_k \leq P_{(m)} \leq \sum_{k=1}^{n} d_k S_k \tag{4.2.6}$$

se sigue a partir de la desigualdad

$$\sum_{k=0}^{n} c_k J_k \leq I(\nu_n \geq m) \leq \sum_{k=0}^{n} d_k J_k. \tag{4.2.7}$$

Entonces, la idea es determinar los coeficientes c_k y d_k para obtener la desigualdad tipo Bonferroni (4.2.6) a partir de la Desigualdad (4.2.7) aplicando técnicas de análisis combinatorio. Éste procedimiento se sigue como consecuencia del siguiente teorema.

Definición 4.1. *Función Booleana.*
Sean $A_1, \ldots, A_n$ conjuntos en la sigma álgebra $\mathcal{F}$. Una función Booleana F de n variables se define como

$$F = f(A_1, \ldots, A_n),$$

donde f opera con los conjuntos $A_1, \ldots, A_n$, de tal manera que F es cualquier combinación de uniones e intersecciónes entre los mismos.

En el Apéndice (A) se presenta el Ejemplo (A.1) de una función Booleana.

Teorema 4.1. *Sean las funciones Booleanas arbitrarias de n variables sobre $\mathcal{F}$:*

$$F_i = f_i(A_1, \ldots, A_n), \quad para \quad i = 1, \ldots, t,$$

y sean $c_1, \ldots, c_t$ *números reales arbitrarios.*

La condición necesaria y suficiente para que se cumpla la desigualdad

$$\sum_k c_k P(F_i) \geq 0 \tag{4.2.8}$$

para cualesquiera elementos $A_1, \ldots, A_n$ *de* $\mathcal{F}$*, es que la Desigualdad* (4.2.8) *se cumpla en los casos especiales donde* $A_1, \ldots, A_n$ *son iguales a* Ω *ó* $\emptyset$.

Demostración.

($\Rightarrow$) Si la Desigualdad (4.2.8) se cumple para cualesquiera $A_1, \ldots, A_n$ en $\mathcal{F}$, entonces se cumple para los casos especiales en los que $A_1, \ldots, A_n$ son iguales a Ω ó $\emptyset$.

($\Leftarrow$) Supongamos que la Desigualdad (4.2.8) se cumple para $A_1^*, \ldots, A_n^*$ en $\mathcal{A}^* = \{\Omega, \emptyset\}$. Consideremos $A_1, \ldots, A_n$ arbitrarias en $\mathcal{F}$. Vamos a demostrar que para cada $\omega \in \Omega$, la desigualdad

$$\sum_k c_k I(F_k)(\omega) \geq 0$$

implica la desigualdad

$$\sum_k c_k P(F_k) \geq 0.$$

Por definición, cada función Booleana de $A_1, \ldots, A_n$ puede ser escrita como una unión de conjuntos del tipo $A_{i_1} A_{i_2} \cdots A_{i_\lambda} A_{i_{\lambda+1}}^c \cdots A_{i_n}^c$, entonces la colección de las $F_i's$ generan una colección $\mathcal{N}$ de subconjuntos de $\{1, 2, \cdots, n\}$, con los subíndices de cada producto de conjuntos. Por ejemplo, un subconjunto sería $\{1, 2, \ldots, \lambda\}$, tal que para cada $\omega \in \Omega$ podemos encontrar una única $J = J(\omega) \in \mathcal{N}$ con:

$$\sum_k c_k I(F_k)(\omega) = C(\omega) \cdot I\left(\prod_{i \in J} A_i \cdot \prod_{i \notin J} A_i^c\right) = C(\omega) \tag{4.2.9}$$

donde $\sum_k^* c_k = C(\omega)$, mientras que $\sum^*$ es sobre todas las k con:

$$\prod_{i \in J} A_i \cdot \prod_{i \notin J} A_i^c \subset \mathcal{F}_k.$$

Observemos que C es una variable aleatoria ya que depende de $J = J(\omega)$.

Por otro lado, haciendo:

$$A_i^* = \begin{cases} \Omega & si \quad i \in J \\ \emptyset & si \quad i \notin J \end{cases}$$

y $F_i^* = f_i^*(A_1^*, \ldots, A_n^*)$, por la Igualdad (4.2.9) tenemos que:

$$\sum_k c_k I(F_k^*)(\omega) = C(\omega) \cdot I\left(\prod_{i \in J} A_i^* \cdot \prod_{i \notin J} A_i^{*c}\right) = C(\omega).$$

Tomando esperanzas por ambos lados de la igualdad anterior se obtiene por hipótesis que:

$$\sum_k c_k P(F_k^*) = E(C) \geq 0.$$

Asimismo, tomando esperanzas por ambos lados de la Igualdad (4.2.9), tenemos que:

$$\sum_k c_k P(F_k) = E(C),$$

entonces se cumple que:

$$\sum_k c_k P(F_k) \geq 0,$$

para cada $\omega \in \Omega$.

■

Observación 4.1. *Una consecuencia inmediata del Teorema (4.1) es que para funciones Booleanas arbitrarias F_k, la desigualdad del tipo*

$$\sum_k c_k P(F_k) \geq 0$$

se obtiene a partir de la desigualdad

$$\sum_k c_k I(F_k) \geq 0.$$

El usar esta implicación para obtener las desigualdades probabilísticas se le conoce como el Método de las Indicadoras.

4.3 Desigualdades tipo Bonferroni

En esta sección demostraremos algunas desigualdades por el método de las indicadoras, y para ello primero introduciremos algunos resultados preliminares.

Denotemos

$$\Delta_{k,s} = \begin{cases} 1 & si \quad k \geq s \\ 0 & si \quad k < s \end{cases}$$

$$\delta_{k,s} = \begin{cases} 1 & si \quad k = s \\ 0 & si \quad k \neq s \end{cases}$$

Lema 4.1. *Para* $k \geq 0$,

$$S_k = E\left[\binom{\nu_n}{k}\right] = \sum_{j=k}^{n} \binom{j}{k} P(\nu_n = j).$$

Demostración.
Para $k = 0$, recordemos que $S_0 = 1$. Por otro lado, se tiene que $\binom{j}{0} = 1$ y por lo tanto

$$\sum_{j=0}^{n} P(\nu_n = j) = P(\nu_n \leq n) = 1 - P(\nu_n > n) = 1,$$

con lo cual se cumple la desigualdad.

Para $k \geq 1$, observemos por un lado que:

$$\sum_{j=k}^{n} \binom{j}{k} P(\nu_n = j) = \sum_{j=k}^{n} \frac{j!}{k!(j-k)!} P(\nu_n = j) = E\left[\binom{\nu_n}{k}\right].$$

Ahora veamos que $J_k = \binom{\nu_n}{k}$.
Para $\nu_n = j$, con $k \leq j \leq n$, tenemos que:

$$\binom{\nu_n}{k} = \frac{j!}{k!(j-k)!}.$$

Por otro lado, si el evento $[\nu_n = j]$ ocurre, entonces ocurren j eventos, por lo cual:

$$\begin{aligned} J_k &= \sum I(A_{i_1}) I(A_{i_2}) \cdots I(A_{i_k}) \\ &= \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} I(A_{i_1} A_{i_2} \cdots A_{i_k}) \\ &= \frac{j!}{k!(j-k)!}. \end{aligned}$$

Por lo anterior tenemos que $J_k = \binom{\nu_n}{k}$ y usando la Igualdad (4.2.5) se llega a que

$$S_k = E\left[\binom{\nu_n}{k}\right].$$

■

Observación 4.2.

S_k se le conoce como el k-ésimo momento binomial de la variable aleatoria ν_n o de los eventos $A_1, \ldots, A_n$.

Lema 4.2. *Para $0 \leq k \leq n-1$ tenemos que*

$$\frac{S_k}{\binom{n}{k}} \geq \frac{S_{k+1}}{\binom{n}{k+1}}$$

Demostración.

Por el Lema (4.1) y el Teorema (4.1) debemos demostrar que para un entero no negativo arbitrario k, con $k \leq n-1$ y $0 \leq \nu_n \leq n$, se cumple:

$$\frac{\binom{\nu_n}{k}}{\binom{n}{k}} \geq \frac{\binom{\nu_n}{k+1}}{\binom{n}{k+1}}.$$

Para ello notemos que

$$\begin{aligned}
\binom{\nu_n}{k+1} &= \frac{\nu_n!}{(k+1)!(\nu_n-(k+1)!)} \\
&= \frac{1}{k+1} \cdot \frac{\nu_n!}{k!(\nu_n-(k+1))!} \\
&= \frac{1}{k+1}\left(\frac{\nu_n-k}{\nu_n-k}\right)\frac{\nu_n!}{k!(\nu_n-(k+1))!} \\
&= \frac{\nu_n-k}{k+1} \cdot \frac{\nu_n!}{k!(\nu_n-k)!} \\
&= \frac{\nu_n-k}{k+1} \cdot \binom{\nu_n}{k},
\end{aligned}$$

y por otro lado,

$$\begin{aligned}
\binom{n}{k+1} &= \frac{n!}{(k+1)!(n-(k+1))!} \\
&= \frac{n-k}{k+1} \cdot \frac{n!}{k!(n-k)(n-(k+1))!} \\
&= \frac{n-k}{k+1} \cdot \binom{n}{k}.
\end{aligned}$$

Por lo tanto, para completar la demostración debemos mostrar que

$$\frac{\binom{\nu_n}{k}}{\binom{n}{k}} \geq \frac{\binom{\nu_n}{k}\frac{\nu_n-k}{k+1}}{\binom{n}{k}\frac{n-k}{k+1}},$$

es decir, que

$$\frac{\frac{\nu_n-k}{k+1}}{\frac{n-k}{k+1}} \leq 1,$$

lo cual se sigue de lo siguiente

$$\frac{\frac{\nu_n-k}{k+1}}{\frac{n-k}{k+1}} = \frac{(\nu_n-k)(k+1)}{(n-k)(k+1)} = \frac{\nu_n-k}{n-k} \leq 1,$$

ya que $\nu_n \leq n$. ■

Desigualdad 4.1. *Para $0 \leq m \leq n$, con $n \geq 1$, y para enteros $d \geq 0$ y $r \geq 0$ se cumple*

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m} \leq P(\nu_n = m) \leq \sum_{k=0}^{2r} (-1)^k \binom{k+m}{m} S_{k+m}, \tag{4.3.1}$$

y

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1} S_{k+m} \leq P(\nu_n \geq m) \leq \sum_{k=0}^{2r} (-1)^k \binom{k+m-1}{m-1} S_{k+m}. \tag{4.3.2}$$

Demostración.

Consideremos la siguiente identidad de análisis combinatorio:

$$\binom{k+m}{m}\binom{a}{k+m} = \binom{a}{m}\binom{a-m}{k} \tag{4.3.3}$$

Por otro lado, demostraremos por inducción sobre b que se cumple la siguiente igualdad:

$$\sum_{k=0}^{b} (-1)^k \binom{a-m}{k} = (-1)^b \binom{a-m-1}{b} \tag{4.3.4}$$

para $a > m$.

Para $b = 1$, por un lado tenemos que

$$\sum_{k=0}^{1} (-1)^k \binom{a-m}{k} = 1 - \binom{a-m}{1} = 1 + m - a,$$

Mientras que por otro lado

$$(-1)^1\binom{a-m-1}{1} = 1+m-a.$$

Ahora suponemos que la igualdad se cumple para $b = n$, es decir,

$$\sum_{k=0}^{n}(-1)^k\binom{a-m}{k} = (-1)^n\binom{a-m-1}{n},$$

y mostraremos que se cumple para $b = n+1$, para lo cual tenemos lo siguiente:

$$\begin{aligned}
\sum_{k=0}^{n+1}(-1)^{n+1}\binom{a-m}{k} &= \sum_{k=0}^{n}(-1)^n\binom{a-m}{k} + (-1)^{n+1}\binom{a-m}{n+1}\\
&= (-1)^n\binom{a-m-1}{n} + (-1)^{n+1}\binom{a-m}{n+1}\\
&= (-1)^n\left[\binom{a-m-1}{n} - \binom{a-m}{n+1}\right]\\
&= (-1)^n\left[-\binom{a-m-1}{n+1}\right]\\
&= (-1)^{n+1}\binom{a-m-1}{n+1}.
\end{aligned}$$

Por lo tanto se cumple la Igualdad (4.3.4).

Combinando las Igualdades (4.3.3) y (4.3.4) se obtiene que:

$$\sum_{k=0}^{b}(-1)^k\binom{k+m}{m}\binom{\nu_n}{k+m} = \begin{cases} (-1)^b\binom{\nu_n}{m}\binom{\nu_n-m-1}{b} & si \quad \nu_n > m\\ 1 & si \quad \nu_n = m\\ 0 & si \quad \nu_n < m \end{cases} \tag{4.3.5}$$

Tomando esperanza en ambos lados de la igualdad anterior y aplicando el Lema (4.1) se sigue que:

$$\sum_{k=0}^{b}(-1)^k\binom{k+m}{m}S_{k+m} = P(\nu_n = m) + (-1)^b\binom{\nu_n-m-1}{b}S_mP(\nu_n > m) \tag{4.3.6}$$

De aquí, para $b = 2r$, esta expresión toma la forma

$$\sum_{k=0}^{2r}(-1)^k\binom{k+m}{m}S_{k+m} = P(\nu_n = m) + (-1)^{2r}\binom{\nu_n-m-1}{2r}S_mP(\nu_n > m),$$

lo cual implica

$$P(\nu_n = m) \leq \sum_{k=0}^{2r} (-1)^k \binom{k+m}{m} S_{k+m}. \tag{4.3.7}$$

Similarmente, para $b = 2d+1$, en la Identidad (4.3.6), obtenemos

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m} + \binom{\nu_n - m - 1}{2d+1} S_m P(\nu_n > m) = P(\nu_n = m),$$

y por lo tanto se cumple

$$P(\nu_n = m) \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m}. \tag{4.3.8}$$

Combinando las Desigualdades (4.3.7) y (4.3.8) se concluye que:

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m} \leq P(\nu_n = m) \leq \sum_{k=0}^{2r} (-1)^k \binom{k+m}{m} S_{k+m},$$

lo cual demuestra la Desigualdad (4.3.1). La Desigualdad (4.3.2) se demuestra aplicando argumentos similares. ∎

Desigualdad 4.2. *Para enteros $0 \leq m < n$, con $n \geq 1$, $d \geq 0$ y $u \geq 0$ se cumple que:*

$$\begin{aligned} P(\nu_n = m) &\geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m} + \frac{2d+2}{n-m} \binom{2d+m+2}{m} S_{2d+m+2} \\ &\leq \sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} S_{k+m} - \frac{2u+1}{n-m} \binom{2u+m+1}{m} S_{2u+m+1} \end{aligned}$$

Demostración.

Primeramente mostraremos que se cumple la cota superior. Para ello, sea λ un número real con el cual se cumple:

$$\delta_{m,t} \leq \sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} \binom{t}{k+m} - \lambda \binom{t}{2u+m+1} \tag{4.3.9}$$

para $t = 0, 1, \ldots, n$.

Si $0 \leq t \leq m + 2u$, entonces

$$\binom{t}{2u+m+1} = 0,$$

por lo cual la Desigualdad (4.3.9) no depende del valor de λ y nos queda:

$$\delta_{m,t} \leq \sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} \binom{t}{k+m}.$$

Entonces, tomando esperanzas en ambos lados de la desigualdad anterior se obtiene:

$$P(\nu_n = m) \leq \sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} S_{k+m},$$

lo cual se cumple por la Desigualdad (4.3.1).

Ahora, para $m + 2u + 1 \leq t \leq n$, aplicando la Igualdad (4.3.6) con $\nu_n = t$, tenemos que:

$$\sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} \binom{t}{k+m} = (-1)^{2u} \binom{t}{m} \binom{t-m-1}{2u} \quad t > m.$$

Luego, considerando de nueva cuenta la Desigualdad (4.3.9) se obtiene que:

$$\delta_{m,t} \leq \binom{t}{m} \binom{t-m-1}{2u} - \lambda \binom{t}{2u+m+1},$$

y como $\delta_{m,t} \geq 0$, se sigue que:

$$\lambda \binom{t}{2u+m+1} \leq \binom{t}{m} \binom{t-m-1}{2u},$$

de aquí:

$$\lambda \leq \frac{\binom{t}{m} \binom{t-m-1}{2u}}{\binom{t}{2u+m+1}} = \frac{2u+1}{t-m} \binom{2u+m+1}{m},$$

por lo cual, el valor más grande que puede tomar λ es

$$\lambda^* = \frac{2u+1}{n-m} \binom{2u+m+1}{m}.$$

Por último, sustituyendo λ^* en la Desigualdad (4.3.9) y luego tomando esperanza por ambos lados nos queda:

$$P(\nu_n = m) \leq \sum_{k=0}^{2u} (-1)^k \binom{k+m}{m} S_{k+m} - \frac{2u+1}{n-m} \binom{2u+m+1}{m} S_{2u+m+1},$$

con lo cual queda demostrada la cota superior.

Ahora, para demostrar la cota inferior de $P(\nu_n = m)$ consideremos de nueva cuenta un número real λ tal que se cumple la siguiente desigualdad:

$$\delta_{m,t} \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} \binom{t}{k,m} + \lambda \binom{t}{2d+m+2} \tag{4.3.10}$$

para $t = 0, \ldots, n$.

Para $0 \leq t \leq 2d+m+1$, tenemos

$$\binom{t}{2d+m+2} = 0,$$

por lo cual la Desigualdad (4.3.10) nos queda:

$$\delta_{m,t} \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} \binom{t}{k+m}.$$

Por lo tanto, tomando esperanzas en ambos lados de la desigualdad anterior, se sigue que

$$P(\nu_n = m) \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} S_{k+m},$$

lo cual se cumple por la Desigualdad (4.3.1).

Por otro lado, para $2d+m+2 \leq t \leq n$, aplicando la Igualdad (4.3.6) con $\nu_n = t$, tenemos que:

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+m}{m} \binom{t}{k+m} = (-1)^{2d+1} \binom{t}{m} \binom{t-m-1}{2d+1} \quad t > m.$$

Sustituyendo lo anterior a la Desigualdad (4.3.10) nos queda:

$$\delta_{m,t} \geq -\binom{t}{m}\binom{t-m-1}{2d+1} + \lambda \binom{t}{2d+m+2} \geq 0,$$

y de aquí

$$\lambda \binom{t}{2d+m+2} \geq \binom{t}{m}\binom{t-m-1}{2d+1},$$

es decir:

$$\lambda \geq \frac{\binom{t}{m}\binom{t-m-1}{2d+1}}{\binom{t}{2d+m+2}} = \frac{2d+2}{t-m}\binom{2d+m+2}{m}.$$

Así, el mínimo valor de λ esta dado por:

$$\lambda^* = \frac{2d+2}{t-m}\binom{2d+m+2}{m},$$

y por lo tanto, sustituyendo el valor de λ^* en la Desigualdad (4.3.10) y tomando esperanzas en ambos lados se obtiene que:

$$P(\nu_n = m) \geq \sum_{k=0}^{n}(-1)^k\binom{k+m}{m}S_{k+m} + \frac{2d+2}{n-m}\binom{2d+m+2}{m}S_{2d+m+2}.$$

Con lo anterior queda terminada la demostraciíon. ■

Para establecer la siguiente desigualdad haremos uso de los siguientes hechos los cuales se encuentran en [5].

Para enteros $n \geq 1$, $m = 0, \ldots, n-1$, $d \geq 0$, y $t = 0, 1, \cdots, n$, denotemos

$$f(2d;m;t) = \sum_{k=0}^{2d}(-1)^k\binom{m+k-1}{m-1}\binom{t}{k+m}. \tag{4.3.11}$$

Además, esta expresión se puede representar como:

$$\begin{aligned} f(2d;m;t) =& 1 + (-1)^{2d}(2d+1)(m+2d+1)\binom{2d+m}{m-1}\binom{t}{m+2d+1} \\ & \int_0^1\int_0^1 x^{2d+m}(1-y)^{2d}(1-xy)^{t-m-2d-1}dxdy. \end{aligned} \tag{4.3.12}$$

De aquí, observemos que:

$$\frac{f(2d;m;t)-1}{\binom{t}{m+2d+1}} = (2d+1)(m+2d+1)\binom{2d+m}{m-1}\int_0^1\int_0^1 x^{2d+m}(1-y)^{2d}(1-xy)^{t-m-2d-1}dxdy. \tag{4.3.13}$$

Notemos que $0 \leq y \leq 1$ y que $0 \leq x \leq 1$, por lo cual $0 \leq (1-xy) \leq 1$. De aquí se tiene

$$0 \leq (1-xy)^{t-m-2d-1} \leq 1.$$

De esta manera, se puede observar que a medida que t crece, la expresión $(1-xy)^{t-m-2d-1}$ decrece, por lo cual el mínimo se alcanza cuando $t=n$.

Desigualdad 4.3. *Para enteros* $0 \leq m < n$, *con* $n \geq 1$, $d \geq 0$, *y* $r \geq 0$ *se cumple*

$$\begin{aligned} P(\nu_n \geq m) &\geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1} S_{k+m} + M(2d+1;m;n) S_{2d+m+2} \\ &\leq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1} S_{k+m} - M(2d;m;n) S_{2d+m+1} \end{aligned}$$

donde

$$M = \frac{(-1)^j \left[\sum_{k=0}^{j} (-1)^k \binom{r+k-1}{r-1}\binom{s}{k+r} - 1\right]}{\binom{s}{r+j+1}}$$

Demostración.

Primeramente demostraremos la cota superior de $P(\nu_n \geq m)$. Para ello, sea β un número real que satisface la desigualdad

$$\Delta_{t,m} \leq \sum_{k=0}^{2d} (-1)^k \binom{m+k-1}{m-1}\binom{t}{k+m} - \beta \binom{t}{2d+m+1} \tag{4.3.14}$$

para $t = 0, \ldots, n$.

Si $0 \leq t \leq 2d+m$, tenemos que $\binom{t}{2d+m+1} = 0$, por lo cual la Desigualdad (4.3.14) se convierte en

$$\Delta_{t,m} \leq f(2d;m;t).$$

Tomando esperanzas en ambos lados de la desigualdad anterior se obtiene

$$P(\nu_n \geq m) \leq \sum_{k=0}^{2d} (-1)^k \binom{m+k-1}{m-1} S_{k+m},$$

lo cual se cumple por la Desigualdad (4.3.1).

Si $2d+m+1 \leq t \leq n$, como $t > m$, de la Desigualdad (4.3.14) y usando la Igualdad (4.3.11) podemos obtener

$$\beta \leq \frac{f(2d;m;t) - 1}{\binom{t}{2d+m+1}}.$$

El argumento de la Igualdad (4.3.13) implica que el valor β más grande que puede tomar está dado por:

$$\beta^* = \frac{f(2d;m;t)-1}{\binom{n}{m+2d+1}},$$

donde

$$M(2d;m;n) = \frac{f(2d;m;t)-1}{\binom{n}{m+2d+1}},$$

con lo cual queda demostrada la cota superior.

Ahora demostraremos la cota inferior. Sea α un número real que cumple la siguiente desigualdad:

$$\Delta_{t,n} \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1}\binom{t}{k+m} + \alpha\binom{t}{2d+2+m}. \tag{4.3.15}$$

Para $0 \leq t \leq 2d+m+1$, se tiene que $\binom{t}{2d+m+2} = 0$, por lo cual la Desigualdad (4.3.15) nos queda:

$$\Delta_{t,m} \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1}\binom{t}{k+m}.$$

Tomando esperanzas en ambos lados de la desigualdad anterior llegamos a que:

$$P(\nu_n \geq m) \geq \sum_{k=0}^{2d+1} (-1)^k \binom{k+m-1}{m-1} S_{k+m},$$

lo cual se cumple por la Desigualdad (4.3.1).

Consideremos ahora el caso cuando $t > 2d+m+1$. Aquí podemos reescribir la Desigualdad (4.3.15) como sigue:

$$1 \geq f(2d+1;m;t) + \alpha\binom{t}{2d+m+2}.$$

Entonces

$$\alpha \leq \frac{1-f(2d+1;m;t)}{\binom{t}{2d+m+2}}. \tag{4.3.16}$$

Recordando la representación para $f(2d+1; m; t)$ dada por la Igualdad (4.3.12) y por la Igualdad (4.3.13), podemos inferir que el lado derecho de la Desigualdad (4.3.16) también decrece a medida que t crece, con lo cual se concluye que el mayor valor para α está dado por:

$$\alpha^* = \frac{1 - f(2d+1; m; n)}{\binom{n}{2d+m+2}},$$

donde $M(2d+1; m; n) = \frac{1-f(2d+1;m;n)}{\binom{n}{2d+m+2}}$.

■

A continuación presentaremos desigualdades para $P_{(m)}$ en términos sólo de S_1 y S_2.

Desigualdad 4.4. *Para* $n \geq m \geq 1$:

$$P(\nu_n \geq m) \leq \begin{cases} \frac{2}{m(m-1)} S_2 & si \quad S_1 \geq \frac{2}{m-1} S_2 \quad y \quad m > 1. \\ \frac{m+n-1}{mn} S_1 - \frac{2}{mn} S_2 & si \quad S_1 < \frac{2}{m-1} S_2 \quad y \quad m > 1. \end{cases}$$

Además las cotas presentes son las mejores entre las cotas del tipo:

$$\{aS_1 + bS_2\} \tag{4.3.17}$$

donde a *y* b *son números reales.*

Demostración.

Sean a y b números reales con la propiedad de que

$$aS_1 + bS_2 \geq P_{(m)}. \tag{4.3.18}$$

Por otro lado, tomemos

$$t = \binom{\nu_n}{1}$$

con lo cual tenemos que $t - 1 = \nu_n - 1$ y

$$\frac{t(t-1)}{2} = \binom{\nu_n}{2}.$$

Por la Observación (4.1) podemos obtener la siguiente desigualdad

$$at + \frac{b}{2}t(t-1) \geq \Delta_{t,m} \tag{4.3.19}$$

para $t = 0, \ldots, n$, la cual es equivalente a la Desigualdad (4.3.18).

Denotaremos el lado izquierdo de la Desigualdad (4.3.19) por $f(t)$:

$$f(t) = at + \frac{b}{2}t(t-1). \tag{4.3.20}$$

Ahora vamos a determinar las constantes a y b con los siguientes casos.

<u>Si $b = 0$:</u>

Bajo esta situación tenemos que $f(t) = at$, y entonces, el mejor valor para a será aquel que sea la menor constante que cumpla:

$$at \geq 1 \quad si \quad t \geq m.$$

Por lo cual, en este caso, la mejor elección es tomar $a = \frac{1}{m}$. Por lo tanto se concluye que para $b = 0$ se tiene $a = \frac{1}{m}$.

<u>Si $a = 0$:</u>

La función dada en (4.3.20) nos queda:

$$f(t) = \frac{b}{2}t(t-1).$$

Luego, el mejor valor para b será aquel que sea la menor constante que cumpla:

$$\frac{b}{2}t(t-1) \geq 1 \quad si \quad t \geq m.$$

Entonces, si $a = 0$ se sigue que $b = \frac{2}{m(m-1)}$.

<u>Si $a, b \neq 0$:</u>

Observemos que para $t = 1$ tenemos que $f(1) = a$, y como $f(t) \geq 0$ entonces se tiene que $a > 0$.

<u>Si $b > 0$:</u>

Tomando $a = \frac{1}{m}$ entonces podemos tomas cualquier valor para $b > 0$, ya que se sigue cumpliendo:

$$at + \frac{b}{2}t(t-1) \geq 1 \qquad si \quad t \geq m.$$

En cambio, si tomamos $b = \frac{2}{m(m-1)}$ entonces podemos tomar cualquier valor $a > 0$, ya que se sigue cumpliendo la desigualdad. Por lo cual, si $b > 0$, podemos tomar $a = \frac{1}{m}$ y $b = \frac{2}{m(m-1)}$.

Si $b < 0$:

Notemos que en este caso $f(t)$ es un polinomio de segundo grado. Especificamente estamos tratando con una parábola con $f(0) = 0$.
Fijando $f(m) = f(n) = 1$ podemos determinar de manera única los coeficientes a y b resolviendo el siguiente sistema de dos por dos:

$$am + \frac{b}{2}m(m-1) = 1$$

$$an + \frac{b}{2}n(n-1) = 1,$$

del cual obtenemos que

$$a = \frac{n+m-1}{mn} \quad y \quad b = -\frac{2}{mn}.$$

Ahora analizaremos los casos anteriores con el fin de encontrar los mejores valores para a y b que hacen una mejor cota en la Desigualdad (4.3.19).
Primeramente veamos que la los valores obtenidos al suponer que $a > 0$ y $b < 0$ corresponden a una mejor cota que los valores obtenidos al suponer que $b = 0$, es decir queremos mostrar que:

$$\frac{1}{m}S_1 > \frac{m+n-1}{mn}S_1 - \frac{2}{mn}S_2. \tag{4.3.21}$$

Para ello vamos a suponer que

$$S_1 < \frac{2}{m-1}S_2 \quad y \quad m > 1,$$

con lo cual tenemos que

$$(m-1)S_1 < S_2,$$

y multiplicando por $(-\frac{1}{mn})$ en ambos lados nos queda

$$-\frac{m-1}{mn}S_1 > -\frac{2}{mn}.$$

Por último sumando en ambos lados la constante $\frac{m+n-1}{mn}S_1$ en la desigualdad anterior, obtiene la Desigualdad (4.3.21) con la condición $S_1 < \frac{2}{m-1}S_2$ y $m > 1$.

Ahora veamos que los valores obtenidos en el caso que $a > 0$ y $b < 0$ son los mejores que cualquier otra cota superior tipo Bonferroni de la forma:

$$a^*S_1 + b^*S_2 \quad con \quad a^* \quad y \quad b^* < 0.$$

Denotemos

$$g(t) = a^*t + b^*\frac{t(t-1)}{2},$$

y observemos que

$$g(0) = 0$$

y

$$g(t) \geq \Delta_{t,m}.$$

Con esto podemos inferir que la menor $g(t)$ que cumpla lo anterior está dada cuando $g(m) = 1$ y $g(n) = 1$, por lo cual tenemos que:

$$g(t) \geq f(t)$$

para $t = 0, 1, \ldots, n$.

Para finalizar la demostración, por último veamos que los valores obtenidos en el caso $a = 0$ son mejores que los obtenidos en el caso $a > 0$ y $b < 0$ con la suposición inicial de que $S_1 \geq \frac{2}{m-1}S_2$ y $m > 1$. Es decir queremos demostrar que

$$\frac{2}{m(m-1)}S_2 \leq \frac{n+m-1}{mn}S_1 - \frac{2}{mn}S_2,$$

siempre que $S_1 \geq \frac{2}{m-1}S_2$ y $m > 1$.

Notemos lo siguiente:

$$\begin{aligned}
\frac{m+n-1}{n}S_1 - \frac{2}{n}S_2 &= (1 + \frac{m-1}{n})S_1 - \frac{2}{n}S_2 \\
&= S_1 + \frac{m-1}{n}S_1 - \frac{2}{n}S_2 \\
&\geq \frac{2}{m-1}S_2 + \frac{m-1}{n}S_1 - \frac{2}{n}S_2.
\end{aligned}$$

Es decir, hasta aquí tenemos que

$$\frac{m+n-1}{n}S_1 - \frac{2}{n}S_2 \geq \frac{2}{m-1}S_2 + \frac{m-1}{n}S_1 - \frac{2}{n}S_2,$$

luego, por hipótesis se tiene que $\frac{m-1}{n}S_1 - \frac{2}{n}S_2 \geq 0$. Entonces se sigue que

$$\frac{m+n-1}{n}S_1 - \frac{2}{n}S_2 \geq \frac{2}{m-1}S_2,$$

con lo cual, multiplicando en ambos lados por $\frac{1}{m}$, se finaliza la demostración. ■

La Desigualdad (4.3.17) establece cotas superiores para $P_{(m)}$, ahora estableceremos cotas inferiores en términos de S_1 y S_2.

Desigualdad 4.5. *Para toda $n \geq 2$ se cumple*

$$P(\nu_n \geq 2) \geq \frac{2}{n(n-1)} S_2,$$

y para $m = 3, 4, \ldots, n$ se satisface:

$$P(\nu_n \geq m) \geq \begin{cases} 0 & si \quad S_1 \geq \frac{2}{m-2} S_2 \\ \\ -\frac{m-2}{n(n+1-m)} S_1 + \frac{2}{n(n+1-m)} S_2 & si \quad S_1 < \frac{2}{m-2} S_2 \end{cases}$$

Además las presentes cotas son las mejores entre las cotas del tipo $\{aS_1 + bS_2\}$, donde a y b son números reales.

Demostración.

Sean las constantes $a, b \in \mathbb{R}$ tal que cumplen la desigualdad:

$$aS_1 + bS_2 \leq P_{(m)}$$

lo cual, por la Observación (4.1), es equivalente a la desigualdad

$$at + \frac{b}{2} t(t-1) \leq \Delta_{t,m} \tag{4.3.22}$$

para $t = 0, 1, \ldots, n$.

A continuación buscaremos valores para a y b que cumplan la Desigualdad (4.3.22).

(i) Si $b = 0$:

La Desigualdad (4.3.22) nos queda como

$$at \leq \Delta_{t,m},$$

con $t = 0, 1, \ldots, n$.

$(i.1)$ Si $t \geq m$, en este caso,

$$at \leq 1,$$

y el mejor valor para a es el mayor valor que cumpla que

$$at \leq 1 \quad con \quad t \geq m.$$

Para lo cual con $a = \frac{1}{n}$, lo anterior se satisface.

$(i.2)$ Si $t < m$,

$$at < 0.$$

Podemos ver, de aquí, que el mayor valor para a que cumpla la desigualdad anterior para $t < m$ es $a = 0$. Sin embargo, con $a = 0$ se cumple que

$$at \leq \Delta_{t,m} \quad con \quad t = 0, 1, \ldots, n.$$

Por lo tanto, si $b = 0$ entonces $a = 0$.

(ii) **Si** $a = 0$: En este caso, la Desigualdad (4.3.22) se convierte en

$$\frac{b}{2}t(t-1) \leq \Delta_{t,m},$$

con $t = 0, 1, \ldots, n$.

$(ii.1)$ Para $t \geq m$, resulta que

$$\frac{b}{2}t(t-1) \leq 1,$$

y el mejor valor para b es aquel que satisface

$$\frac{b}{2}t(t-1) \leq 1 \quad con \quad t \geq m.$$

Tomando $b = \frac{2}{n(n-1)}$ se cumple lo anterior, y , más aún, se cumple para toda $t = 0, 1, \ldots, n$.

$(ii.2)$ Para $t < m$, obtenemos de la Desigualdad (4.3.22)

$$\frac{b}{2}t(t-1) \leq 0.$$

Si $m = 2$, entonces $t = 0, 1$, y así $t(t-1) = 0$. Con ello se puede tomar cualquier valor para b, ya que se seguirá cumpliendo la desigualdad. En particular podemos tomar $b = \frac{2}{n(n-1)}$. Si $m > 2$, entonces $t(t-1) \neq 0$, por lo cual el mayor valor para b que cumpla la desigualdad es $b = 0$.

Como resultado tenemos dos casos:

- Para $m = 2$ tenemos que con $b = \frac{2}{n(n-1)}$ se cumple la desigualdad deseada.
- Para $m > 2$ y con $b = 0$ también se satisface la desigualdad deseada.

(iii) Si $a, b \neq 0$:
Sustituyendo $t = 1$ en la Desigualdad (4.3.22) se obtiene que $a \leq 0$.

$(iii.1)$ Para $b < 0$, notemos que en este caso se puede tomar cualquier valor para $b < 0$, ya que se seguirá cumpliendo la Desigualdad (4.3.22).

$(iii.2)$ Para $b > 0$, definamos

$$f(t) = at + \frac{b}{2}t(t-1),$$

y notemos que $f(t)$ es una parábola con $f(0) = 0$. Luego, poniendo $f(m-1) = 0$ y $f(n) = 1$ podemos determinar de manera única los valores de a y b resolviendo el sistema:

$$a(m-1) + \frac{b}{2}(m-1)(m-2) = 0$$

$$an + \frac{b}{2}n(n-1) = 1.$$

Así:

$$a = -\frac{m-2}{n(n+1-m)} \quad y \quad b = \frac{2}{n(n+1-m)}$$

satisfacen la Desigualdad (4.3.22).

Ahora analizaremos cuáles son las mejores cotas.

(I) Veamos que la cota obtenida en (iii) es mejor que la obtenida en (i) suponiendo que $S_1 < \frac{2}{m-2}S_2$ y $m > 2$, es decir demostraremos que

$$-\frac{m-2}{n(n+1-m)} + \frac{2}{n(n+1-m)} > 0.$$

Para esto, de la hipótesis resulta que:

$$(m-2)S_1 \leq 2S_2,$$

luego, multiplicando en ambos lados por $\frac{1}{n(n+1-m)}$ se sigue la desigualdad deseada.

(II) La cota obtenida en (ii) es mejor que la cota obtenida en (iii) cuando $S_1 > \frac{2}{m-2}S_2$ y $m > 2$, en otras palabras veamos que:

$$-\frac{(m-2)}{n(n+1-m)} + \frac{2}{n(n+1-m)}S_2 < 0,$$

lo cual se sigue inmediatamente de la suposición que $S_1 > \frac{2}{m-2}S_2$.

(III) La cota obtenida en $(iii.2)$ es la mejor cota inferior del tipo:

$$a^*S_1 + b^*S_2,$$

con $a^* < 0$ y $b^* > 0$.

Para demostrarlo consideremos $g(t) = a^*t + b^*t(t-1)$, del cual notemos que $g(0) = 0$, y

$g(t)$ debe cumplir que $g(t) \leq \Delta_{t,m}$ para $t = 0, 1, \ldots, n$.
Con esta observación podemos inferir que la máxima función que cumpla lo anterior está dada cuando $g(m-1) = 0$ y $g(n) = 1$, por lo cual

$$g(t) \leq f(t),$$

para $t = 0, 1, \ldots, n$.

Con el analisis anterior queda demostrada la desigualdad. ■

Desigualdad 4.6. *Para $r \geq 1$ y $2r \leq k \leq n$ tenemos que*

$$P(\nu_n \geq 1) \geq \sum_{i=1}^{2r} (-1)^{i+1} \binom{2r}{i} \frac{S_i}{\binom{k}{i}}. \tag{4.3.23}$$

Demostración.
Consideremos

$$g(t) = \sum_{i=1}^{2r} (-1)^{i+1} \binom{2r}{i} \frac{\binom{t}{i}}{\binom{k}{i}}. \tag{4.3.24}$$

Necesitamos mostrar que

$$g(t) \leq \Delta_{t,1} \quad t = 0, 1, \ldots, n. \tag{4.3.25}$$

Como $g(0) = 0$, entonces se cumple la desigualdad para $t = 0$. Ahora vamos a demostrar que se satisface para $t \geq 1$. Para ello observemos lo siguiente:

$$
\begin{aligned}
g(t) &= 1 - \sum_{i=1}^{2r} (-1)^i \binom{2r}{i} \frac{\binom{t}{i}}{\binom{k}{i}} \\
&= 1 - \frac{1}{\binom{k}{2r}} \sum_{i=1}^{2r} (-1)^i \binom{t}{i} \frac{\binom{2r}{i}\binom{k}{2r}}{\binom{k}{i}} \\
&= 1 - \frac{1}{\binom{k}{2r}} \sum_{i=1}^{2r} (-1)^i \binom{t}{i} \binom{k-i}{k-2r} \\
&= 1 - \frac{1}{\binom{k}{2r}} \binom{k-t}{2r}, \quad \textit{por} \quad [7] \\
&= 1 - \frac{1}{\binom{k}{2r}} \left[\frac{(k-t)(k-t-1)\cdots(k-t-2r+1)}{2r!} \right] \\
&= 1 - \frac{1}{\binom{k}{2r}} \left[\frac{(t-k)(t-k-1)\cdots(t-k-2r+1)}{2r!} \right]
\end{aligned}
$$

- Si $0 < t < k - 2r + 1$, entonces $t - k - 1 + 2r < 0$, y por lo tanto tenemos:

$$\left[\frac{(t-k)(t-k-1)\cdots(t-k-2r+1)}{2r!} \right] > 0,$$

ya que contiene un numero par de términos negativos.

- Si $k - 2r + 1 \leq t \leq k$, entonces:

$$\left[\frac{(t-k)(t-k-1)\cdots(t-k-2r+1)}{2r!} \right] = 0.$$

- Si $t > k$, el término dentro del corchete es positivo.

Esto muestra que para $t \geq 1$ se tiene $g(t) \leq 1$, es decir se cumple la Desigualdad (4.3.25), y aplicandole a la misma la Observación (4.1) se completa la demostración. ■

Desigualdad 4.7. *Para enteros k y r, con $2r \leq k \leq n$, se satisface que:*

$$P(\nu_n \geq 1) \leq S_1 - \sum_{i=2}^{2r} (-1)^i \binom{2r}{i} \frac{ik - (2r+1)(i-1)}{2r - i + 1} \cdot \frac{S_i}{\binom{k}{i}} + \frac{2r+1}{\binom{k}{2r}} S_{2r+1}$$

Demostración.
Consideremos la equivalencia de la cota inferior probada en la Desigualdad (4.6):

$$\sum_{i=1}^{2r} (-1)^{i+1} \frac{\binom{2r}{i}\binom{t}{i}}{\binom{k}{i}} \leq 1 \qquad t = 1, 2, \ldots, n.$$

Luego, si multiplicamos en ambos lados de la desigualdad por $(1-t)$ y enseguida le sumamos a ambos lados t se obtiene:

$$t - (-t) \sum_{i=1}^{2r} (-1)^{i+1} \frac{\binom{2r}{i}\binom{t}{i}}{\binom{k}{i}} \geq (1-t) + t,$$

es decir

$$t + (1-t) \sum_{j=1}^{2r} (-1)^{i+1} \frac{\binom{2r}{i}\binom{t}{i}}{\binom{k}{i}} \geq 1, \tag{4.3.26}$$

para $t = 1 \ldots, n$. De aquí, aplicando la identidad:

$$t\binom{t}{i} = (i+1)\binom{t}{i+1} + i\binom{t}{i}$$

podemos llegar a que la Desigualdad (4.3.26) es equivalente a:

$$t - \sum_{i=2}^{2r} (-1)^i \binom{2r}{i} \frac{ik - (2r+1)(i-1)}{2r - i + 1} \frac{\binom{t}{i}}{\binom{k}{i}} + \frac{2r+1}{\binom{k}{2r}} \binom{t}{2r+1} \geq \Delta_{t,1}$$

para $t = 1, \ldots, n$. Sin embargo observemos que también se sigue cumpliendo para $t = 0$.

De esta manera, aplicando la Observación (4.1) a la última desigualdad se finaliza la demostración. ■

4.4 Aplicaciones

Teoremas Límite

El siguiente teorema es una aplicación de la Desigualdad (4.3.1), ya que es una herramienta fundamental en la demostración de dicho teorema.

Teorema 4.2. *Sean* $A_{1,N}, A_{2,N}, \ldots, A_{n,N}$, $\quad n = n(N), N \geq 1$, *sucesiones de eventos sobre un espacio de probabilidad* $(\Omega_N, \mathcal{A}_N, P_N)$, *donde* $n = n(N) \to \infty$ *cuando* $N \to \infty$. *Sea* $\nu_{n,N}$ *la variable aleatoria que mide el número de eventos* $A'_{j,N}s$ *que ocurren y sea* $S_{k,N}$ *el* k*-ésimo momento binomial de* $\nu_{n,N}$. *Supongamos que*

$$\lim_{N\to\infty} S_{k,N} = a(k) < \infty, \quad para \quad k \geq 0 \quad fijo, \tag{4.4.1}$$

y que

$$p(r) = \sum_{k=0}^{\infty} (-1)^k \binom{k+r}{r} a(k+r), \quad r \geq 0, entero, \tag{4.4.2}$$

converge. Entonces

$$\lim_{N\to\infty} P_N(\nu_{n,N} = r) = p(r), \quad r \geq 0.$$

Demostración.
Inicialmente tenemos que

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+r}{r} S_{k+r,N} \leq P_N(\nu_{n,N} = r) \leq \sum_{k=0}^{2d} (-1)^k \binom{k+r}{r} S_{k+r,N}$$

con $d \geq 0$ fijo. Ahora, tomemos limite cuando $N \to \infty$ en cada desigualdad de la expresión anterior:

$$\lim_{N\to\infty} \sum_{k=0}^{2d+1} (-1)^k \binom{k+r}{r} S_{k+r,N} \leq \lim_{N\to\infty} P_N(\nu_{n,N} = r) \leq \lim_{N\to\infty} \sum_{k=0}^{2d} (-1)^k \binom{k+r}{r} S_{k+r,N}$$

y por hipótesis, lo anterior nos queda como sigue

$$\sum_{k=0}^{2d+1} (-1)^k \binom{k+r}{r} a(k+r) \leq \lim_{N\to\infty} P_N(\nu_{n,N} = r) \leq \sum_{k=0}^{2d} (-1)^k \binom{k+r}{r} a(k+r), \tag{4.4.3}$$

lo cual se obtiene para cada $d \geq 0$ fijo.

En seguido, hagamos $d \to \infty$ en la expresión (4.4.3), y por hipótesis obtenemos lo siguiente

$$p(r) \leq \lim_{N\to\infty} P_N(\nu_{n,N} = r) \leq p(r),$$

por lo tanto,

$$\lim_{N\to\infty} P_N(\nu_{n,N} = r) = p(r).$$

■

Podemos observar, este teorema establece convergencia en probabilidad de la variable aleatoria $\nu_{n,N}$ en el límite de la sucesión de las familias dadas.

Corolario 4.1. *Teorema de la Convergencia de Poisson.*
Si en el Teorema (4.2) tomamos en particular $a(k) = \frac{\lambda^k}{k!}$, *con algún* $0 < \lambda < \infty$, *entonces*

$$\lim_{N\to\infty} P_N(\nu_{n,N} = r) = \frac{\lambda^r e^{-\lambda}}{r!}, \quad r \geq 0.$$

Demostración.
Como $a(k) = \frac{\lambda^k}{k!}$, en nuestro caso, $p(r)$ tiene la siguiente forma

$$\begin{aligned} p(r) &= \sum_{k=0}^{\infty} (-1)^k \binom{k+r}{r} \frac{\lambda^{k+r}}{(k+r)!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{(k+r)!}{r!k!} \cdot \frac{\lambda^{k+r}}{(k+r)!} \\ &= \frac{\lambda^r}{r!} \sum_{k=0}^{\infty} (-1)^k \frac{\lambda^k}{k!} \\ &= \frac{\lambda^r}{r!} e^{-\lambda}. \end{aligned}$$

Por último, sustituyendo el valor de $p(r)$ en el resultado del teorema (4.2) se obtiene que

$$\lim_{N\to\infty} P(\nu_{n,N} = r) = \frac{\lambda^r}{r!} e^{-\lambda}.$$

■

El resultado del corolario nos expresa que podemos obtener la distribución de Poisson como un caso particular de la convergencia en probabilidad de la variable aleatoria $\nu_{n,N}$ en el límite de la sucesión de las familias dadas. Lo cual es natural suponerlo por como esta definida la variable aleatoria $\nu_{n,N}$, y además concuerda con la aproximación de la distribución de Poisson por medio de la distribución Binomial.

Construcción de intervalos de confianza para parámetros desconocidos

A continuación se presenta con un ejemplo general el uso de las desigualdades tipo Bonferroni para la construcción de intervalos de confianza para cantidades desconocidas.

Sea $\theta = (\theta_1, \ldots, \theta_k)$ un parámetro k-dimensional desconocido. Estamos interesados en establecer intervalos de confianza para cada valor θ_i de θ, para lo cual podemos definir a cada evento A_i como el intervalo donde no se encuentra el valor de θ_i. De esta manera, $P_{[0]} = P(\nu_k(A) = 0)$ proporciona cotas para la probabilidad de que cada intervalo establecido en A_i esté el valor del parámetro θ_i. Notemos que la más simple aproximación para $P_{[0]}$ esta dada por

$$P(\nu_k(A) = 0) \geq 1 - S_1.$$

Problema de conteo

Se tiene una caja con n cartas etiquetadas con los números $1, \ldots, n$. Las cartas son repartidas aleatoriamente una por una y sin remplazo. Cada repartición tiene la misma probabilidad de ser seleccionada, esto implica que cuando todas las n cartas son repartidas, hay $n!$ posibilidades de reparticiones distintas y cada una de éstas tiene una probabilidad de $\frac{1}{n!}$ para ser seleccionada.

Diremos que una coincidencia ocurre en la $j-$ésima repartición si la carta escogida está marcada con j. Calcularemos la probabilidad de que no haya coincidencias en la elección de las cartas.

Para resolver el problema, sea A_j el evento que representa que en la $j-$ésima repartida hay coincidencia. En estos términos nos interesa estimar el valor de

$$P_{[0]} = P(\nu_n = 0).$$

Notemos que para algún $1 \leq i_1 \leq i_2 \cdots i_k \leq n$ se tiene que:

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = \frac{(n-k)!}{n!}.$$

Luego,

$$S_k = \sum \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!}$$

En este caso S_k representa la probabilidad de todas las posibilidades en las que se puede encontrar k coincidencias en una repartición.

Ahora, aplicando la Desigualdad (4.3.1) para $m = 0$ se obtiene:

$$\sum_{k=0}^{2d+1} (-1)^k S_k \leq P(\nu_n = 0) \leq \sum_{k=0}^{2r} (-1)^k S_k.$$

Para estimar estos valores, observemos que tenemos:

$$1 - S_1 + S_2 - S_3 + \cdots + (-1)^u S_u = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^u \frac{1}{u!},$$

lo cual es una cota superior o inferior de $P_{[0]}$, dependiendo si u es par o impar, con $u \leq n$. De aquí, si hacemos tender $n \to \infty$, y por consiguiente $u \to \infty$, obtenemos:

$$\lim_{n \to \infty} P(\nu_n = 0) = e^{-1}.$$

Notemos que e^{-1} es una buena aproximación a $P_{[0]}$ para valores pequeños de n, considerando que tenemos la siguiente expresión:

$$\left| P(\nu_n = 0) - e^{-1} \right| \leq \frac{1}{(n+1)!}.$$

Capítulo 5

Cotas para Probabilidades de Colas

5.1 Introducción

Sea X una variable aleatoria con distribución normal $N(\mu, \sigma^2)$. Es bien conocido (ver [2]) que esta distribución tiene la siguiente propiedad:

$$P[|X-\mu| > k\sigma] \leq \frac{2\phi(k)}{k}, \quad k > 0, \tag{5.1.1}$$

donde $\phi(k)$ es la densidad de una variable aleatoria normal estándar.

En la Desigualdad (5.1.1) se observa que la función de distribución y la densidad están relacionadas de tal manera que se pueden calcular cotas para la probabilidad de la cola de la distribución de una variable aleatoria normal en términos de su densidad.

El objetivo del presente capítulo es introducir desigualdades de este tipo para una clase más amplia de variables aleatorias. Específicamente, daremos condiciones bajo las cuales la cota para la probabilidad de la cola de la variable aleatoria puede expresarse en términos de su densidad.

Este tipo de desigualdades son consecuencia de un resultado general el cual es establecido en la Sección (5.2). Después, en la Sección (5.3) presentamos ejemplos de desigualdades para distribuciones específicas.

5.2 Cotas en términos de la densidad

Sea X una variable aleatoria con función de densidad $f(x)$. En esta sección mostraremos que bajo condiciones apropiadas existe una constante $D(k) = D(k, f)$, con $k \geq 0$, tal que

$$P[|X| > x] \leq D(k)f(k).$$

Proposición 5.1. *Sea X una variable aleatoria con función de distribución $F(x)$, $|X|$ tiene una función de densidad $f(x)$, y $E|X|^r < \infty$ para algún $r > 0$. Suponga que f es diferenciable y que $\psi(x) := -\frac{xf'(x)}{f(x)}$ es creciente para $x > 0$. Entonces para cualquier k tal que $\psi(k) > 1$*

$$P(|X| > k) \leq \frac{k}{\psi(k) - 1} f(k). \tag{5.2.1}$$

Demostración.
Definamos

$$G(x) = x^r P(|X| > x),$$

donde $x > 0$ y $r > 0$. Observemos que la función G satisface las siguientes propiedades:

(a). $G(0) = 0$.

(b). G toma la forma:

$$\begin{aligned} G(x) &= x^r P(|X| > x) \\ &= x^r [1 - P(|X| \leq x)] \\ &= x^r - x^r F_{|X|}(x) \end{aligned}$$

(c). Denotando $Y = |X|$ tenemos,

$$\begin{aligned} EY^r &= \int_0^\infty y^r dF_Y(y) \\ &\geq \int_x^\infty y^r dF_Y(y) \\ &\geq \int_x^\infty x^r dF_Y(y) \\ &= x^r P(|X| > x) \\ &= G(x), \end{aligned}$$

para $x > 0$. Con lo cual $G(x) \leq E|X|^r$.

(d). Por el punto (c), se obtiene

$$G(x) \leq \int_x^\infty y^r dF_Y(y) \leq EY^r < \infty,$$

y como

$$\lim_{x\to\infty} \int_x^\infty y^r dF_Y(y) = 0,$$

entonces se concluye que

$$\lim_{x\to\infty} G(x) = 0.$$

Por los puntos (a)-(d) concluimos que $G(x)$ tiene un máximo en $[0,\infty)$. Para encontrar el máximo se calculará la derivada de G, usando que $F'_{|X|}(x) = f(x)$, como sigue

$$\begin{aligned} G'(x) &= rx^{r-1} - rx^{r-1}F_{|X|}(x) - x^r f(x) \\ &= rx^{r-1}\left(1 - F_{|X|}(x)\right) - x^r f(x) \\ &= rx^{r-1}P\left(|X| > x\right) - x^r f(x) \end{aligned}$$

Haciendo $G'(x) = 0$ se obtiene que el $\sup_{x>0} G(x)$ se alcanza en el punto $x_* = x_*(F,r)$ tal que

$$P\left(|X| > x_*\right) = \frac{x_* f(x_*)}{r},$$

por lo cual

$$\sup_{x>0} G(x) = \sup_{x>0} \{x^r P\left(|X| > x\right)\} = \frac{x_*^{r+1} f(x_*)}{r}.$$

Además, como

$$\frac{x_*^{r+1} f(x_*)}{r} \leq \frac{\sup_{x>0}\left\{x^{r+1} f(x)\right\}}{r},$$

obtenemos

$$P\left(|X| > y\right) \leq \frac{\sup_{x>0}\left\{x^{r+1} f(x)\right\}}{ry^r}, \tag{5.2.2}$$

con $y > 0$ y $r > 0$. Ahora, en particular tomemos $r = \psi(k) - 1$, con k tal que $\psi(k) > 1$, y definamos $J(x) = x^{r+1} f(x)$. Derivando con respecto a x, se llega a que

$$\begin{aligned} J'(x) &= (r+1)x^r f(x) + x^{r+1} f'(x) \\ &= x^r\left[(r+1)f(x) + xf'(x)\right], \end{aligned}$$

y, de nueva cuenta, haciendo $J'(x) = 0$, la función J se maximiza cuando

$$r = -\frac{xf'(x)}{f(x)} - 1 = \psi(x) - 1.$$

Sin embargo, usando el hecho que $r = \psi(k) - 1$, tenemos $\psi(x) = \psi(k)$, y como ψ es creciente para $x > 0$ se concluye que $k = x$. Por último, sustituyendo lo anterior en la Desigualdad (5.2.2) nos queda

$$P(|X| > y) \leq \frac{k^{r+1} f(k)}{r y^r},$$

y en particular si $y = k$ se tiene que

$$P(|X| > k) \leq \frac{k^{r+1} f(k)}{r k^r} = \frac{k f(k)}{\psi(k) - 1},$$

lo cual demuestra la propisición. ■

5.3 Desigualdades para probabilidades de colas

Desigualdad 5.1. *Sea X una variable aleatoria con densidad t de Student $f_\alpha(x)$ y con α grados de libertad. Entonces para todo $k > 1$,*

$$P(X > k) \leq \frac{k(\alpha + k^2)}{\alpha(k^2 - 1)} f_\alpha(k). \tag{5.3.1}$$

Demostración.
Primero, por definición tenemos que

$$f_\alpha(x) = \frac{\Gamma(\frac{\alpha+1}{2}) \left(1 + \frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}}}{\sqrt{\alpha\pi}\,\Gamma(\frac{\alpha}{2})}$$

para $x \geq 0$, y así

$$f_\alpha'(x) = -\frac{\Gamma(\frac{\alpha+1}{2}) \frac{\alpha+1}{2} \left(1 + \frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}-1} \frac{2x}{\alpha}}{\sqrt{\alpha\pi}\,\Gamma(\frac{\alpha}{2})}.$$

Con lo anterior podemos calcular

$$\psi(x) = \frac{x\Gamma(\frac{\alpha+1}{2})\frac{(\alpha)}{2}\left(1+\frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}-1}\frac{2x}{\alpha}}{\Gamma(\frac{\alpha+1}{2})\left(1+\frac{x^2}{\alpha}\right)^{-\frac{\alpha+1}{2}}}$$

$$= \frac{x\left(\frac{\alpha+1}{2}\right)\left(\frac{2x}{\alpha}\right)}{1+\frac{x^2}{\alpha}}$$

$$= \frac{x^2(\alpha+1)}{\alpha+x^2}.$$

Observemos que $\psi(x)$ es creciente, y además

$$\psi(x) - 1 = \frac{\alpha(x^2-1)}{\alpha+x^2}.$$

Entonces, para $k > 1$ se tiene que $\psi(k) > 1$. Por lo tanto, se cumplen las condiciones de la Proposición (5.1) y de esta manera obtenemos

$$P(X > k) \leq \frac{k(\alpha+k^2)}{\alpha(k^2-1)} f_\alpha(k),$$

para todo $k > 1$. ■

Desigualdad 5.2. *Sea X una variable aleatoria con densidad Gamma*

$$f_{\alpha,\lambda}(x) = e^{-}x^{\alpha-1}\frac{\lambda^\alpha}{\Gamma(\alpha)}, \qquad x > 0.$$

Entonces para todo $k > E(X) = \frac{\alpha}{\lambda}$

$$P(X > k) \leq \frac{k}{k\lambda - \alpha} f_{\alpha,\lambda}(x). \tag{5.3.2}$$

Demostración.

Derivando $f_{\alpha,\lambda}(x)$ obtenemos

$$f'_{\alpha,\lambda}(x) = e^{-\lambda x}x^{\alpha-2}\frac{\lambda^\alpha}{\Gamma(\alpha)(-\lambda x+\alpha-1)}.$$

De aquí,

$$\psi(x) = \frac{-xe^{-\lambda x}x^{\alpha-2}\frac{\lambda^2}{\Gamma(\alpha)}() - \lambda x + \alpha - 1}{e^{-\lambda x}x^{\alpha-1}\frac{\lambda^\alpha}{\Gamma(\alpha)}}$$

$$= \lambda x - \alpha + 1.$$

Es facil ver que $\psi(x)$ es creciente, y para $k > \frac{\alpha}{\lambda}$ se tiene que $\psi(x) > 1$. Por lo tanto, por la Proposición (5.1), obtenemos la desigualdad deseada. ■

En general, la técnica para obtener este tipo de desigualdades es partir de la densidad correspondiente y verificar si cumple las condiciones de la Proposición (5.1). A continuación presentamos dos desigualdades más, para la distribución F y Beta, las cuales se obtienen aplicando argumentos similares a las desigualdades anteriormente obtenidas.

Desigualdad 5.3. *Sea X una variable aleatoria con densidad F, denotada por $f_{\alpha,\beta}(x)$, con α y β grados de libertad. Entonces para $k > 1$*

$$P(X > k) \leq \frac{2k(\beta + \alpha k)}{\alpha\beta(k-1)} f_{\alpha,\beta}(k). \tag{5.3.3}$$

Desigualdad 5.4. *Sea X una variable aleatoria con densidad $Beta(\alpha, \beta)$, $f_{\alpha,\beta}$. Si $\beta > 1$, entonces para todo $k > \frac{\alpha}{\alpha+\beta-1}$*

$$P(X > k) \leq \frac{k(1-k)}{k(\alpha + \beta - 1) - \alpha} f_{\alpha,\beta}(k). \tag{5.3.4}$$

Apéndice

Capítulo A

En el presente apéndice se establecen resultados sin demostración que se usan para el desarrollo de la tesis. En algunos de estos resultados se hace referencia de libros en los cuales se encuentra su demostración.

Definición A.1. *Variable Aleatoria Simétrica.*
Una variable aleatoria X y su distribución son simétricas si X y $-X$ tienen la misma distribución.

Definición A.2. *Variable Aleatoria Simetrizada.*
Sea X una variable aleatoria con función característica $f(t)$. Consideremos la variable aleatoria $X^s = X - Y$, donde Y es una variable aleatoria independiente de X y tiene la misma distribución que X. La variable aleatoria X^s es llamada la variable aleatoria simetrizada.

Lema A.1. *Sea X una variable aleatoria. Sea $a, b \in \mathbb{R}, 0 < q < 1$.*

(i) Si $P(X \leq a) > q$ entonces cada cuantil k_q de la variable aleatoria X satisface la condición $k_q \leq a$.

(ii) Si $P(X \geq b) > q$ entonces cada cuantil k_{1-q} satisface la condición $k_{1-q} \geq b$.

(iii) Si $P(X \leq a) \geq q$ entonces existe un cuantil k_q de la variable aleatoria X que satisface la condición $k_q \leq a$.

(iv) Si $P(X \leq b) \geq q$ entonces existe un cuantil k_{1-q} tal que $k_{1-q} \geq b$.

Este Lema es establecido en [8].

Desigualdad A.1. *Desigualdad de Chebyshev.*
Sea X una variable aleatoria con esperanza finita μ. Entonces, para todo $k > 0$,

$$P(|X-\mu| \geq k) \leq \frac{Var(X)}{k^2}.$$

Desigualdad A.2. *Desigualdad de Markov.*
Si W es una variable aleatoria no negativa, entonces

$$P[W \geq r] \leq \frac{E[W^r]}{r^k},$$

para toda $k, r > 0$.

Lema A.2. *Sea $g(x)$ una función no negativa, no decreciente en la parte positiva de la recta real y tal que $g(0) = 0$. Sea X una variable aleatoria albitraria. Si $Eg(X) < \infty$, entonces*

$$Eg(X) = \int_0^\infty P(|X| \geq x) dg(x).$$

Este Lema es demostrado en [8].

Desigualdad A.3. *Desigualdad de Lyapunov.*
Sea Y una variable aleatoria albitraria y sea $0 < r < s$, entonces

$$(E|Y|^r)^{\frac{1}{r}} \leq (E|Y|^s)^{\frac{1}{s}}.$$

Teorema A.1. *Sea la variable aleatoria X, $\beta_r = E|X|^r$ y $0 < r < s$. Luego se cumple que*

$$\beta_r^{\frac{1}{r}} \leq \gamma^{\frac{1}{r}-\frac{1}{s}} \beta_s^{\frac{1}{s}},$$

donde $\gamma = P(X \neq 0)$.

Este se demuestra en [8].

Ejemplo A.1. *Función Booleana.*
Consideremos el mismo espacio de probabilida $(\Omega, \mathcal{F}, P)$ definido en el Capítulo (4). Sean $A_1, A_2, A_3 \in \mathcal{F}$, definamos

$$F = f(A_1, A_2, A_3) = A_1 A_2 \cup A_3^c,$$