Leji 10

Universidad Nacional Autónoma de México

Facultad de Ciencias

"ANALISIS DEL TRABAJO DE EMILE BOREL SOBRE SUCESIONES INFINITAS DE ENSAYOS BERNOULLI"

T E S I S

Que para obtener el Título de

MATEMATICO

p r e s e n t a

MIGUEL ANGEL GUTIERREZ ANDRADE

MEXICO, D. F. 1983

CONTENIDO

Página

Introducción I

1 Probabilidades Numerables

1.1 Caso independiente. 5
1.2 Análisis del caso independiente. 14
1.3 Caso dependiente. 23
1.4 Análisis del caso dependiente. 30
1.5 Lema de Borel-Cantelli 34

2 Las Fracciones Continuadas

2.1 Generalidades. 36
2.2 Aplicación de Borel de las probabilidades numerables a las fracciones continuadas. 47
2.3 Análisis de la aplicación de Borel a las fracciones continuadas. 56
2.4 Crítica de Felix Bernstein a los resultados obtenidos por E. Borel acerca de fracciones continuadas. 58
2.5 Respuesta de Borel a la crítica de Bernstein. 68

3 Extensión del Lema de Borel-Cantelli

3.1 Esperanza condicional. 70
3.2 Martingalas. 78
3.3 Lema de Borel-Cantelli extendido. 80
3.4 Aplicación de la extensión del Lema de Borel-Cantelli a sucesiones infinitas de ensayos Bernoulli. 83

Apéndice

Referencias

INTRODUCCION

En teoría de probabilidad estudiamos teoremas límites los cuales son de dos tipos distintos: teoremas límite fuertes (Aquellos que tratan con la convergencia casi dondequiera de una sucesión de variables aleatorias) y teoremas límite débiles (Aquellos que tratan con la convergencia en probabilidad de una sucesión de variables aleatorias.

Aquí tratamos el surgimiento de un teorema límite fuerte; el Lema de Borel-Cantelli, relacionado con la ley fuerte de los grandes números. Empezaremos por establecer las leyes de los grandes números.

El teorema de Bernoulli establece que para un número suficientemente grande de ensayos n, la probabilidad de la desigualdad

$$\left| \frac{S_n}{n} - p \right| < \varepsilon$$

es mayor que $1 - \eta$ para un arbitrario $\eta > 0$, donde S_n es el número de casos favorables en los n ensayos y p es la probabilidad del caso favorable. O en otras palabras, cuando incrementamos n, la probabilidad que el número promedio de casos favorables se desvíe de p más de una ε previamente asignada, tiende a cero. En 1909 el matemático francés Emile Borel detectó una proposición más profunda la cual es conocida como la ley fuerte de los grandes números. Este resultado puede establecerse para el caso de ensayos Bernoulli como sigue

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = p \right) = 1$$

la cual mejora esencialmente al teorema de Bernoulli (Ley débil de los grandes números). Asegura la existencia de una frecuencia límite igual a la probabilidad teórica p, para todos los posibles "experimentos" excepto posiblemente un conjunto de ellos con probabilidad cero (Pero no necesariamente un conjunto vacío).

Este trabajo pretende rescatar los conceptos y pruebas originales dados por E. Borel del Lema de Borel-Cantelli y su aplicación a las fracciones continuadas, así como hacer un análisis de los mismos. Cabe mencionar aquí, que las secciones (1.1), (1.3), (2.2) y (2.4), están escritas, salvo algunas extensiones, como originalmente fueron escritas.

En el capítulo 1, se presenta el Lema de Borel-Cantelli para sucesiones infinitas de ensayos Bernoulli (A este proceso infinito, Borel le llamó "Probabilidades Numerables"). Originalmente el resultado fue probado para ensayos independientes (1909), pero Borel aplica este resultado a las fracciones continuadas, definiendo sus ensayos de tal manera que no resultan independientes; esta dependencia la hace notar F. Bernstein en 1911 al obtener un resultado de las fracciones continuadas (Sección 2.4) que él cree que está en contradicción con los ya obtenidos por Borel. Borel responde, en 1912, con este caso dependiente (Sección 1.3) haciendo notar que esta corrección no invalida sus resultados originales. Al final de este primer capítulo se enuncia y demuestra el Lema de Borel-Cantelli como actualmente se conoce.

En el capítulo 2 se expone, por un lado, la aplicación del Lema de Borel-Cantelli a las fracciones continuadas (dada por Borel) y por otro lado, los resultados obtenidos por F. Bernstein sobre fracciones continuadas. Bernstein llega a estos resultados, a diferencia de Borel, usando explícitamente probabilidad condicional y las propiedades básicas de las fraccio nes continuadas. Aunque, esencialmente, los resultados de Bernstein concuerdan, en el límite, con los de Borel, Bernstein cree que son contradictorios y lo señala en su artículo de 1911. Finalmente, en 1912, Borel contesta a Bernstein que, en efecto, los ensayos definidos para las fracciones continuadas eran dependientes, pero los resultados obtenidos seguían siendo válidos, Haciendo algunas modificaciones en sus cálculos originales, o podían obtenerse de los mismos cálculos de Bernstein.

En el capítulo 3 se da la extensión al Lema de Borel-Cantelli, haciendo uso de la esperanza condicional y de la teoría de martingalas. Por último, se usa esta extensión para nuevamente demostrar el teorema de las fracciones continuadas de Borel.

CAPITULO 1

Probabilidades Numerables

En 1933 A.N. Kolmogorov publica un trabajo fundamental en la teoría de la probabilidad bajo el título de "*Grundbeg·iffe der Wahrscheinlich einlichkeitsrechnung*" (Fundamentos de Probabi· .dad) que inicia una nueva era en la teoría de probabilidad y sus métodos. En este trabajo, Kolmogorov desarrolla un modelo (conjunto de axiomas) que puede ser asociado a situaciones (Esto es, un experimento, observaciones, etc.) en la que entran factores aleatorios. El modelo de Kolmogorov puede describirse brevemente como sigue: Sea (Ω, L, P) una terna en donde Ω es un espacio abstracto (el espacio de eventos elementales), L es un σ-campo de subconjuntos de Ω (el conjunto de eventos) y $P(E)$ es una medida (la probabilidad del evento E) definida para todo $E \in L$ que satisface las condiciones: $P(\Omega)=1$ y si $A_1, A_2, \ldots$ es una colección numerable de eventos en L tales que $A_i \cap A_j = \phi$ para $i \neq j$ se cumple que

$$P\left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} P(A_n)$$

esta última propiedad es llamada la aditividad numerable.

Antes de esta axiomatización de la teoría de la probabilidad, aparece en 1909 un artículo de Emile Borel bajo el título: "*Les Probabilités Denombrables et Leurs Applications Arithmetiques*" en cuyo primer capítulo trata lo que Borel llama Probabilidades Numerables que más adelante analizaremos. Obviamente, Borel no emplea los resultados de este modelo de Kolmogorov para desarrollar este capítulo, sino que usa algunos principios de probabilidad clásica que a continuación enunciaremos como él los define en su libro "*Eléments de la Théorie des Probabilités*." Para probabilidades discontinuas (caso finito), Borel da la siguiente definición:

Definición de Probabilidad. Se dice que la probabilidad es la razón del número de casos favorables entre el número de casos posibles, cuando todos los casos son mirados como igualmente probables.

Teorema (Probabilidades Totales). Cuando el evento, cuya probabilidad buscada, se puede dar de varias maneras diferentes que son excluyentes recíprocamente, la probabilidad buscada es igual a la suma de las probabilidades parciales correspondientes a estas diversas maneras.

El teorema anterior es lo que Borel llama la Ley de Probabilidades Totales que actualmente llamamos Regla de la Suma.

Teorema (Probabilidades Compuestas). Cuando el evento cuya probabilidad buscada consiste de la produccíon sucesiva de dos eventos, la probabilidad buscada es igual al producto de la probabilidad del primero de estos eventos por la probabilidad para que el segundo se presente cuando el primero se presentó. Más general, si la producción sucesiva de varios eventos es necesaria, se debe multiplicar las probabilidades diversas de estos eventos, evaluando cada uno de ellos dentro de la hipótesis de que los precedentes se hayan presentado.

Es claro que si la probabilidad del segundo no varía cuando el primero se presenta, esta última restricción es inútil; se puede dentro de este caso, considerar los eventos como simultáneos.

Este último teorema, es lo que Borel llama Ley de Probabilidades Compuestas, el último párrafo se refiere a eventos independientes.(Regla del Producto) Borel no usa explícitamente el término independiente.

Por último daremos la definición de Probabilidad Geométrica.

Definición. La probabilidad para que el punto M se encuentre en un cierto segmento PQ de AB, es proporcional a la longitud de este segmento.

Si suponemos que M se encuentra dentro del segmento AB, la probabilidad para que M este en AB es 1; por lo tanto, la probabilidad para que M se encuentre sobre PQ es igual a la razón $\frac{PQ}{AB}$.

En estos principios básicos de probabilidad clásica, Borel apoya los resultados obtenidos en las probabilidades numerables. Borel comienza su artículo de 1909 con las siguientes palabras:

> "Uno distingue generalmente, dentro de los problemas de probabilidad, dos categorías principales, acordando ya sea que el número de casos posibles es finito o infinito: la primera categoría constituye lo que se llama las *probabilidades discontinuas* o probabilidades dentro de un dominio discontinuo, mientras que la segunda categoría comprende las *probabilidades continuas* o *probabilidades geométricas*. Tal clasificación aparece incompleta cuando uno se refiere a los resultados adquiridos en la teoría de conjuntos; entre la cardinalidad de

los conjuntos finitos y la cardinalidad del continuo se encuentra la cardinalidad de los conjuntos numerables; me propongo mostrar brevemente el interés que une las cuestiones de probabilidad en cuyas proposiciones tales conjuntos intervienen; yo los llamaré, para abreviar, *probabilidades numerables*.

Antes de definir más precisamente las probabilidades numerables, me gustaría indicar en pocas palabras los razonamientos en contra de futuras fallas en su estudio. La principal de estas razones es la importancia de la noción de conjunto numerable; esta importancia no ha sido puesta en duda por ningún matemático; pero me parece que es mayor todavía de lo que se piensa.

Muchos analistas, en efecto, ponen en primer lugar la noción del continuo; este concepto interviene de una manera más o menos explícita en sus razonamientos. He indicado recientemente como esta noción del continuo, considerando que tiene una cardinalidad mayor que la numerable, me parece que es una noción puramente negativa. La cardinalidad de los conjuntos numerables solamente podemos conocerla de una forma positiva, únicamente lo último interviene efectivamente en nuestro razonamiento. Es claro, realmente, que el conjunto de elementos analíticos que actualmente se pueden definir y considerar no es otra cosa que un conjunto numerable; yo creo este punto de vista prevalecerá más y más cada día entre los matemáticos y que el continuo probará haber sido un instrumento transitorio, cuya utilidad presente no es ignorada (daremos ejemplos posteriormente), pero vendrá a ser mirado sólamente como un medio de estudio de los conjuntos numerables, los cuales constituyen la única realidad que somos capaces de lograr." Borel(1909: 247-248).

En este punto, Borel divide los problemas de probabilidades numerables en tres categorías:

1° El número de resultados posibles es finito; se hace una infinidad numerable de ensayos.

2° El número de resultados posibles es numerable; se hace un número limitado de ensayos.

3° El número de resultados posibles y de ensayos es numerable.

Para propósitos de este trabajo, únicamente nos interesará los problemas de primera categoría. En 1.1 y 1.3 se repetirán los resultados obtenidos por Borel usando su notación y argumentos.

En estos problemas de primera categoría, Borel considera dos casos; cuando los ensayos son independientes y cuando los ensayos son dependiente. El caso dependiente, históricamente, surgió tres años después (1912) que el caso independiente, debido a una observación hecha por F. Bernstein en 1911 al artículo de Borel de 1909. Esta observación se expondrá en la sección 2.3.

1.1 Caso independiente.

Supondremos primero que el número de casos se reduce a dos. Llamaremos a uno caso favorable y al otro desfavorable. Sean p_1, $p_2, \ldots, p_n, \ldots$ las probabilidades sucesivas del caso favorable dentro de los ensayos de rango 1, 2,..., n,... supondremos todas las p_i, $i=1, 2, \ldots$ diferentes de 0 o de 1 y distinguiremos los casos donde la serie

(1.1) $$p_1 + p_2 + \cdots + p_n + \cdots$$

es convergente o divergente.

a. La serie (1.1) es convergente.

Denotemos por $A_{0,n}$ a la probabilidad de que el caso favorable no se presente en los n primeros ensayos, esta probabilidad está dada, según el principio de probabilidades compuestas,

$$A_{0,n} = (1-p_1)(1-p_2) \cdots (1-p_n)$$

como la serie (1.1) es convergente, este producto tiene un límite cuando n aumenta indefinidamente y este límite está entre cero y uno (véase apéndice). Por lo tanto, la probabilidad que el caso favorable nunca se presente es

(1.2) $$A_0 = \prod_{i=1}^{\infty} (1-p_i)$$

El argumento que Borel da, para justificar este límite, es el siguiente:

"Dentro del caso de convergencia, la extensión del principio de las probabilidades compuestas tenderá por sí, a la definición de la probabilidad buscada. Si se consideran en efecto las n primeras pruebas, los principios clásicos permiten definir y calcular la probabilidad para que el caso favorable no se presente dentro de estas n pruebas: se comprueba que, para n creciente, no solamente esta probabilidad varía muy poco de una manera absoluta, sino que varía también muy poco de manera relativa, es decir que sus variaciones son una fracción muy débil de su valor. Se puede por lo tanto, habiendo asignado tal valor que se vea hasta la precisión relativa que se desee alcanzar, es cierto que esta precisión es efectivamente alcanzada al cabo de un número de pruebas, - puede ser muy grande, mas asignable; el pasaje al límite que hemos efectuado no levanta por lo tanto ninguna dificultad y está entéramente justificada." Borel(1909: 249).

Igual que con $A_{0,n}$ busquemos ahora la probabilidad $A_{1,n}$ para que el caso favorable se presente una sóla vez en los primeros n ensayos. Esto puede ocurrir de n formas diferentes, dependiendo si el caso favorable se presenta en la primera, en la segunda,..., en la n-ésima prueba. Aplicando el principio de probabilidades totales, la probabilidad buscada es,

$$A_{1,n} = p_1(1-p_2) \cdots (1-p_n) + (1-p_1)p_2 \cdots (1-p_n)$$

$$+ \cdots +(1-p_1)(1-p_2) \cdots (1-p_{n-1})p_n$$

o

$$A_{1,n} = (1-p_1)(1-p_2) \cdots (1-p_n) \sum_{i=1}^{n} \frac{p_i}{1-p_i}$$

o también

(1.3) $$A_{1,n} = A_{0,n}(u_1 + u_2 + \cdots + u_n)$$

tomando

$$u_i = \frac{p_i}{1-p_i}$$

Tomando el límite cuando n tiende a infinito en (1.3) obtenemos

(1.4) $$A_1 = A_0(u_1 + u_2 + \cdots + u_n + \cdots)$$

ya que la serie que multiplica a A_0 es convergente puesto que a partir de cierta n tenemos que

$$1-p_n \geq \frac{1}{2}$$

y por lo tanto

$$\sum_{i=n}^{\infty} \frac{p_i}{1-p_i} \leq \sum_{i=n}^{\infty} 2p_i = 2\sum_{i=n}^{\infty} p_i < \infty$$

Así que, la probabilidad para que el caso favorable se presente una y sólamente una vez es

(1.5) $$A_1 = A_0 \sum_{n=1}^{\infty} u_n$$

Borel justifica la relación (1.5) diciendo:

" Un razonamiento análogo al que hemos hecho dentro del pro blema I permite aplicar, dentro del caso de convergencia, el principio de probabilidades totales;... " Borel(1909: 250).

Donde por problema I se refiere al cálculo de A_0.

De igual forma, la probabilidad para que el caso favorable se presente exactamente k veces, dentro de las primeras n pruebas es:

$$A_{k,n} = A_{0,n} \Sigma u_{n_1} u_{n_2} \cdots u_{n_k}$$

donde $n_1, n_2, \ldots, n_k$ son k números distintos tomados de 1, 2,..., n y cada combinación de k índices diferentes figura una y sólamente una vez dentro de la suma. Observemos que

$$\Sigma u_{n_1} u_{n_2} \cdots u_{n_k} < (1+u_1)(1+u_2) \cdots (1+u_n)$$

este último producto tiene un límite cuando n aumenta indefinidamente porque la serie Σu_i es convergente.(véase apéndice) Cualquiera que sea el número k, la probabilidad para que el caso favorable se presente exactamente k veces es:

$$A_k = A_0 \Sigma u_{n_1} u_{n_2} \cdots u_{n_k} \tag{1.6}$$

con $n_1, n_2, \ldots, n_k$ enteros diferentes y cada combinación de índices figura una y sólamente una vez dentro de la suma infinita.

Consideremos la serie

(1.7) $$A_0 + A_1 + A_2 + \cdots + A_k + \cdots$$

Esta serie tiene como suma

$$S = A_0 \left[1 + \Sigma u_i + \Sigma u_{n_1} u_{n_2} + \cdots + \Sigma u_{n_1} u_{n_2} \cdots u_{n_k} + \cdots \right]$$

que se puede poner como

(1.8) $$S = A_0 (1+u_1)(1+u_2) \cdots (1+u_n) \cdots$$

Por otro lado, tenemos

$$1+u_n = 1 + \frac{p_n}{1-p_n} = \frac{1}{1-p_n}$$

y

$$\prod_{n=1}^{\infty} (1+u_n) = \frac{1}{\prod_{n=1}^{\infty} (1-p_n)} = \frac{1}{A_0},$$

Finalmente obtenemos que

$$S = A_0 \frac{1}{A_0} = 1.$$

Si denotamos por A_∞ la probabilidad de obtener una infinidad de casos favorables, podemos decir que la probabilidad A_∞ es inferior a

$$1-S_n$$

para cualquier n, tomando $S_n = A_0 + A_1 + \cdots + A_n$. S_n tiende a 1 a medida que n aumenta, por lo tanto, podemos escoger n de manera que $1-S_n$ sea arbitrariamente pequeña, con lo que se concluye que

$$A_\infty = 0.$$

b. La serie (1.1) es divergente.

La probabilidad para que el caso favorable no se presente en los n primeros ensayos es igual a:

$$A_{0,n} = (1-p_1)(1-p_2) \cdots (1-p_n)$$

y tiende hacia cero cuando n aumenta indefinidamente.(véase apéndice) Por lo tanto $A_0 = 0$. Sin embargo, (en contraposición con el caso convergente en el que Borel acepta que A_0 tiene exáctamente el valor dado por (1.2)) Borel hace la siguiente aclaración con respecto a la interpretación de A_0, para este caso divergente:

" Existe, en efecto, una verdadera discontinuidad entre una probabilidad infinitamente pequeña, es decir una probabilidad variable que tiende hacia cero, y una probabilidad igual a cero.... No existe inconveniente en decir que, dentro del caso de la divergencia, A_0 es nulo; mas no se deberá perder de vista que este lenguaje no significa más que lo siguiente: la probabilidad para que el caso favorable no se presente tiende hacia cero cuando el número de pruebas aumenta infinitamente." Borel(1909: 249)

La probabilidad para que el caso favorable se presente una y sólo una vez es el límite cuando n tiende a infinito de la ecuación (1.3), $A_{0,n}$ tiende hacia cero y la serie $\Sigma\, u_i$ es divergente porque,

$$\sum_{i=1}^{n} p_i < \sum_{i=1}^{n} \frac{p_i}{1-p_i} = \sum_{i=1}^{n} u_i \ .$$

La expresión $A_{1,n}$ se presenta por lo tanto de forma indeterminada. Por otra parte

$$1-p_i < e^{-p_i}$$

Así

$$(1.9) \qquad A_{1,n} < e^{-(p_1 + p_2 + \cdots + p_n)} (u_1 + u_2 + \cdots + u_n)$$

y se ve inmediatamente que la divergencia de la serie (1.1) arrastra como consecuencia que $A_{1,n}$ tienda hacia cero cuando n aumenta indefinidamente.* Por lo tanto

$$A_1 = 0.$$

Se ve igualmente que la probabilidad A_k para que el caso favorable se presente k veces es nula.

Inmediatamente después; Borel observa que, si cada una de las probabi

* Esta afirmación, hecha por Borel, no es correcta en general. Ver página 18.

lidades A_k es nula, la suma S también lo es y en consecuencia A_∞ es igual a la unidad, pero como S es una suma infinita de probabilidades nulas, Borel comenta que:

" El resultado es exacto pero el razonamiento carece de rigor. Es claro, por otra parte, que en este caso no se puede buscar la probabilidad para que el caso favorable se presente una infinidad de veces en n pruebas y luego hacer crecer n indefinidamente; se razonará, por lo tanto, como sigue: escojamos un número fijo m y busquemos la probabilidad para que el caso favorable se presente más de m veces en n ensayos."

esto es

$$1 - (A_{0,n}+A_{1,n}+ \cdots +A_{m,n}) \qquad n > m$$

el límite cuando n aumenta indefinidamente de esta expresión es igual a la unidad para cualquier número fijo m.

" Esto significa que se puede apostar, con ventaja, una suma tan grande como se quiera contra 1 franco, a favor de que el número de casos favorables será superior a un número fijo m cualquiera; esto es precisamente el significado de este enunciado: la probabilidad A_∞ es igual a 1."* Borel(1909: 251).

Se puede resumir así, los resultados obtenidos dentro de este estudio:

---

* Este razonamiento que concluye con que $A_\infty = 1$, es esencialmente lo mismo que hacer A_∞ = 1-S. Ver página 21.

Dentro del caso donde la serie (1.1) es convergente, las probabilidades A_0, A_1,... tienen valores determinados no nulos y la probabilidad A_∞ es nula; dentro del caso de divergencia, las probabilidades A_k son todas nulas y A_∞ es igual a la unidad.

Hasta aquí, hemos supuesto que el número de resultados posibles es igual a dos. Suponemos ahora, que el número de resultados es igual a r. Sean p_i^1, p_i^2,..., p_i^r las probabilidades respectivas de los resultados 1, 2,...,r dentro del ensayo de rango i. Se cumple que

$$p_i^1 + p_i^2 + \dots + p_i^r = 1 ,$$

es claro que la serie $\sum_{i=1}^{\infty} p_i^\rho$ no es convergente para todos los valores de ρ (ρ = 1,...,r). Supongamos que es convergente para ρ = 1, 2,...,h (h<r). Podemos aplicar a cada uno de los casos 1, 2,...,h los resultados que hemos obtenido anteriormente. La probabilidad para que un caso se presente exactamente k veces tiene un valor bien determinado, la probabilidad para que se presente una infinidad de veces es nula. Estos resultados siguen siendo válidos si se mira como caso favorable a cualquier subconjunto de estos h casos. Al contrario, la probabilidad para que un caso de índice mayor que h, se presente un número finito cualquiera de veces es nula y la probabilidad para que se presente una infinidad de veces es igual a la unidad.

I.2 Análisis del caso independiente.

Borel obtiene los resultados anteriores usando la siguiente técnica: si desea saber la probabilidad de que un evento B ocurra en un número infinito de ensayos, primero calcula la probabilidad que el evento B ocurra en n ensayos y luego toma el límite de esta probabilidad cuando n tiende a infinito. Al tomar este límite existen dos casos posibles; la probabilidad del evento B está entre cero y uno o la probabilidad de B es igual a cero o uno. Si la probabilidad de B está entre cero y uno, Borel argumenta que no hay ninguna dificultad en tomar este límite como la probabilidad de que B ocurra, ya que, no existe ninguna "discontinuidad" entre el valor de este límite y los valores que se obtienen al ir tomando n cada vez más grande. No sucede lo mismo para el caso en que la probabilidad de B es cero o uno; Borel analiza cuando la probabilidad de B es cero, (en caso que la probabilidad de B sea igual a uno, Borel analiza la probabilidad del complemento de B) diciendo que existe una "verdadera discontinuidad" entre una probabilidad variable que tiende hacia cero, cuando el número de ensayos crece indefinidamente y una probabilidad igual a cero. Así, la probabilidad de B deberá verse como cero. Borel dice, en este punto, que este resultado no significa de ningún modo la imposibilidad de que el evento B ocurra, sino que, solamente la probabilidad de que B ocurra es tanto más cercana a cero, cuando el número de ensayos es muy grande.

Lo anteriormente dicho, podemos notarlo, volviendo al cálculo de las probabilidades A_0, A_1 y A_∞ . La probabilidad que no se tenga ningún caso favorable en un número infinito de ensayos es

$$A_0 = (1-p_1)(1-p_2) \cdots (1-p_n) \cdots$$

que es el límite cuando n tiende a infinito de

$$A_{0,n} = (1-p_1)(1-p_2) \cdots (1-p_n)$$

Cuando (1.1) es convergente, Borel concluye que, no hay ninguna dificultad en tomar A_0 como este producto infinito ya que

$$0 < A_0 < 1$$

Cuando (1.1) diverge, entonces A_0 deberá verse como igual a cero, pero sin olvidar que es el límite cuando el número de ensayos tiende a infinito.

La probabilidad de que se tenga un caso favorable en el primer ensayo y nunca más vuelva a ocurrir es

$$\bar{\omega}_1 = p_1(1-p_2) \cdots (1-p_n) \cdots$$

y en forma general, la probabilidad que se tenga un caso favorable exclusivamente en el n-ésimo ensayo y los demás casos sean desfavorables en un número infinito de ensayos es

$$\bar{\omega}_n = (1-p_1) \cdots (1-p_{n-1})p_n(1-p_{n+1}) \cdots$$

Por lo tanto, la probabilidad que se tenga un caso favorable en una infinidad de ensayos es

$$\begin{aligned} A_1 &= \bar{\omega}_1 + \bar{\omega}_2 + \cdots + \bar{\omega}_n + \cdots \\ &= A_0(u_1 + u_2 + \cdots) \end{aligned}$$

donde

$$u_n = \frac{p_n}{1-p_n}$$

Esta expresión nuevamente es el límite cuando n tiende a infinito de

$$A_{1,n} = A_{0,n}(u_1 + u_2 + \cdots + u_n)$$

Así que, si (1.1) converge entonces, como ya se vio en la sección 1.1, la serie infinita de término general u_n converge y por lo tanto, A_1 tiene un valor entre cero y uno.

Cuando la serie (1.1) es divergente, la situación se torna muy interesante. Por un lado, la probabilidad A_1 puede ser mirada como la suma infinita de las $\bar{\omega}_n$, siendo $\bar{\omega}_n$ igual a cero para toda n, por lo tanto,

$$A_1 = \sum_{n=1}^{\infty} \bar{\omega}_n = 0$$

pero $\bar{\omega}_n$ es un límite de una probabilidad que tiende a cero, así que, (según Borel) hay una "discontinuidad" al tomar este límite y subraya este hecho diciendo:

" Desde otro punto de vista, se puede decir que la probabilidad buscada A_1 es la suma de las probabilidades $\bar{\omega}_n$ que son todas nulas; mas, su número no siendo finito, se puede concluir sin precaución que la probabilidad total es nula, si se recuerda que la probabilidad nula no significa imposibilidad." Borel(1909: 250).

Así que, aunque la expresión obtenida para A_1 es un límite de una expresión para el caso finito, Borel no emplea la extensión de la ley de probabilidades totales cuando suma probabilidades nulas. Esto queda clāramente establecido en la siguiente cita:

" ..., la probabilidad total que corresponde a una infinidad de probabilidades nulas puede muy bien ser diferente de cero. Si suponemos por ejemplo que existe un medio de escoger dentro de la sucesión indefinida de los enteros un número completamente al azar, de manera que todos los números tengan la misma probabilidad, esta probabilidad deberá forzozamente ser nula; y la probabilidad total será igual a la unidad." Borel(1926: 18).

Borel, nuevamente aquí, usa la técnica descrita para calcular A_1 en el caso divergente, toma un argumento alternativo para mostrar $A_1 = 0$. Si $A_{1,n}$ es la probabilidad que haya exáctamente un éxito en los primeros n ensayos, por (1.9) Borel asegura que

$$\lim_{n\to\infty} A_{1,n} = 0$$

esta conclusión es cierta, pero no se sigue de (1.9).

Se puede dar un contraejemplo tomando:

$$p_n = \frac{e^{2n}}{1+e^{2n}} \quad , \quad q_n = \frac{1}{1+e^{2n}} \quad , \quad u_n = \frac{p_n}{q_n} = e^{2n}$$

Así, $\sum_{n=1}^{\infty} p_n$ claramente es divergente y

$$\sum_{k=1}^{n} u_k > u_n = e^{2n}$$

mientras que

$$\sum_{k=1}^{n} p_k < n$$

de modo que

$$e^{-(p_1 + p_2 + \cdots + p_n)} (u_1 + u_2 + \cdots + u_n) > e^{-n}e^{2n} = e^n$$

esto muestra que no necesariamente

$$e^{-(p_1 + p_2 + \cdots + p_n)} (u_1 + u_2 + \cdots + u_n)$$

tiende a cero cuando la serie (1.1) diverge.

En 1926, Borel corrige este error haciendo la hipótesis adicional de que $q_i = 1-p_i$ sea mayor que un número fijo ε para toda i. Así, $\frac{1}{1-p_i}$ es inferior a $M = \frac{1}{\varepsilon}$, y por lo tanto se cumple

$$u_1 + u_2 + \cdots + u_n < M (p_1 + p_2 + \cdots + p_n)$$

Por otra parte

$$1-p_i < e^{-p_i}$$

por consiguiente

$$A_{1,n} < M\,(p_1 + p_2 + \cdots + p_n)\; e^{-(p_1 + p_2 + \cdots + p_n)}$$

y como la expresión $S_n e^{-S_n}$ tiende a cero cuando S_n aumenta indefinidamente, entonces se cumple que

$$\lim_{n \to \infty} A_{1,n} = A_1 = 0.$$

Obviamente $A_1 = 0$, generalizando el concepto de probabilidad total; pero esta generalización, para Borel, carece de rigor.

Los casos A_2 ,..., A_k,... se tratan similarmente. Para el cálculo de A_∞ , si (1.1) es convergente, entonces Borel se fija en las sumas parciales

$$S_n = \sum_{k=0}^{n} A_k$$

y dice que A_∞ es inferior a $1-S_n$. Tomando el límite cuando n tiende a infinito obtiene

$$A_\infty = 0$$

ya que S_n tiende a uno cuando n tiende a infinito.

Si la serie (1.1) diverge, entonces Borel, no puede dar el mismo argumento ya que $S_n = 0$ para toda n. Tampoco puede usar su técnica de buscar esta probabilidad en un número finito de casos porque no tiene sentido preguntarse por la ocurrencia de un número infinito de casos favorables en n ensayos. Borel justifica que $A_\infty = 1$ de una manera esencialmente equivalente a tomar

$$A_\infty = \lim_{n \to \infty} 1-Sn = 1$$

como en el caso convergente. El argumento intuitivo alternativo que Borel da, es el siguiente: fija un número m y se pregunta por la probabilidad que el caso favorable se presente más de m veces en n ensayos: esta probabilidad está dada por

$$1 - (A_{0,n} + A_{1,n} + \dots + A_{m,n}) \qquad n > m$$

haciendo tender n a infinito, obtiene que esta probabilidad es 1, ya que

$$\lim_{n \to \infty} A_{k,n} = A_k = 0 \qquad \text{para } k = 0, 1, \dots, m$$

Como m puede ser tomada tan grande como se quiera, entonces Borel concluye que $A_\infty = 1$.

Para ver que este argumento es esencialmente el dado en el caso convergente, denotaremos por B_m^n al evento de obtener más de m casos favorables en n ensayos, por B_m el evento de obtener más de m casos favorables en un número infinito de ensayos y por B_∞ el evento de obtener una infi nidad de casos favorables en una infinidad de ensayos.

Borel con el argumento de arriba muestra que

$$\lim_{n \to \infty} P(B_m^n) = 1$$

Y como B_m^n es una sucesión creciente tal que

$$B_m = \bigcup_{n=m+1}^{\infty} B_m^n$$

entonces

$$\lim_{n \to \infty} P(B_m^n) = P(\lim_{n \to \infty} B_m^n) = P(B_m) = 1$$

Sin embargo, este razonamiento no implica que

$$P(B_\infty) = 1$$

El hecho que la probabilidad anterior sea igual a 1 es una consecuencia de la aditividad numerable ya que la sucesión B_m es decreciente y

$$B_\infty = \bigcap_{m=1}^{\infty} B_m$$

Así,

$$\lim_{m \to \infty} P(\bigcap_{k=1}^{m} B_k) = P(\lim_{m \to \infty} \bigcap_{k=1}^{m} B_k) = P(B_\infty)$$

Por lo tanto

$$P(B_\infty) = 1$$

o alternativamente

$$B_\infty^c = \bigcup_{m=1}^{\infty} B_m^c$$

$$P(B_\infty^c) = P\left(\bigcup_{m=1}^{\infty} B_m^c\right)$$

$$\leq \sum_{m=1}^{\infty} P(B_m^c)$$

$$= \sum_{m=1}^{\infty} S_m$$

Como $S_m = 0$ para toda m, se sigue nuevamente que

$$P(B_\infty^c) = 0$$

y por lo tanto

$$P(B_\infty) = 1 .$$

I.3 Caso dependiente.

Borel desarrolla el caso dependiente en su artículo de 1912, como una respuesta a la objeción que F. Bernstein hace de su aplicación de las probabilidades numerables a las fracciones continuadas. Esta discusión se desarrollará en detalle en el siguiente capítulo. Hemos querido presentar aquí el caso dependiente, para comparar inmediatamente sus similitudes y diferencias con el caso independiente. Borel modifica su caso independiente a este que se presentará a continuación con el único fin de poderlo aplicar a los resultados obtenidos en las fracciones continuadas.

Consideremos una infinidad numerable de ensayos sucesivos. Sea p_n la probabilidad del caso favorable para la prueba de rango n, q_n la probabilidad complementaria; p_n y q_n son las probabilidades globales, es decir, aquellas que se obtienen suponiendo que se ignoran los resultados de los n-1 primeros ensayos. Cuando estos ensayos presentan alguna particularidad (por ejemplo, los n-1 primeros ensayos han sido favorables, la probabilidad relativa al n-ésimo ensayo toma en general un valor distinto a p_n. Supondremos que este valor está comprendido entre dos límites conocidos p_n' y p_n'' diferentes, comprendidos entre 0 y 1, para cualquiera que sea n.

Estaremos dentro del caso de convergencia si ambas series

$$\sum_n p_n' \quad , \quad \sum_n p_n''$$

son convergentes; dentro del caso divergente si ambas series son divergentes. Si la primera es convergente y la segunda divergente, se estará dentro de un caso dudoso que dejaremos a un lado.

Mostraremos que para el caso de convergencia, todos los resultados que hemos establecido anteriormente, suponiendo las probabilidades independientes subsisten. De igual forma, veremos que dentro del caso de divergencia se seguirán los mismos resultados obtenidos, suponiendo las probabilidades independientes.

Usaremos para este caso la misma notación que empleamos en el caso in dependiente, designaremos por A_0 la probabilidad para que el caso favora ble no se presente jamás, por A_k la probabilidad para que se presente k veces exactamente, por A_∞ la probabilidad para que se presente una infini dad de veces. Para demostrar que

$$A_\infty = 0$$

o

$$A_\infty = 1 - (A_0 + A_1 + \dots + A_k + \dots)$$

Es suficiente mostrar que la serie entre paréntesis tiene como suma la unidad, es decir, se puede escoger un número k de manera que se cumpla

$$S_k = A_0 + A_1 + \dots + A_{k-1} > 1-\varepsilon$$

siendo ε arbitrariamente pequeña.

Denotemos $p_0 = q_1$ y por p_{11} la probabilidad para que el caso favora ble se presente en el segundo ensayo dado que se presentó en el primero, por p_{10} la probabilidad para que no se presente, con la misma hipótesis en el primer resultado; p_{01} y p_{00} designarán las probabilidades de caso favorable y desfavorable en el segundo ensayo cuando el caso desfavora-- ble se presentó en el primer ensayo. En forma similar se puede generali zar la notación para los ensayos posteriores.

Se tiene evidentemente que

$$p_1 p_{11} + p_0 p_{01} = p_2$$

$$p_1p_{10} + p_0p_{00} = q_2$$

y de la misma manera

$$p_1p_{11}p_{111} + p_1p_{10}p_{101} + p_0p_{01}p_{011} + p_0p_{00}p_{001} = p_3$$

$$p_1p_{11}p_{110} + p_1p_{10}p_{100} + p_0p_{01}p_{010} + p_0p_{00}p_{000} = q_3$$

y así sucesivamente.

Además, se cumple necesariamente

$$p_{11} + p_{10} = 1$$

y de manera general

$$p_{\alpha\beta\gamma\,\cdots\,\lambda 0} + p_{\alpha\beta\gamma\,\cdots\,\lambda 1} = 1 ,$$

donde $\alpha, \beta, \gamma, \ldots, \lambda$ designan los índices 0 o 1.

Por hipótesis se cumple que

$$p_n^{'} \leq p_{\alpha\beta\gamma\,\cdots\,\lambda 1} \leq p_n^{''}$$

para cada valor de n, no importando el valor que puedan tomar los índices $\alpha, \beta, \gamma, \quad , \lambda$.

Calculemos las probabilidades A_k para k=0, 1, 2,... tenemos que

$$A_0 = p_0 p_{00} p_{000} \cdots ,$$

$$A_1 = p_1 p_{10} p_{100} p_{1000} \cdots + p_0 p_{01} p_{010} p_{0100} \cdots$$
$$+ p_0 p_{00} p_{001} p_{0010} \cdots + \cdots + \cdots$$

$$A_2 = p_1 p_{11} p_{110} p_{1100} \cdots + p_1 p_{10} p_{101} p_{1010} \cdots$$
$$+ p_0 p_{01} p_{011} p_{0110} \cdots + p_0 p_{01} p_{010} p_{0101} \cdots + \cdots$$

Designemos por Π_4 , Π_4' ,... los productos de la forma

$$\Pi_4 = p_{0000} p_{00000} p_{000000} \cdots$$

y de manera general,

$$\Pi_4^{(k)} = p_{\alpha\beta\gamma 0} p_{\alpha\beta\gamma 00} p_{\alpha\beta\gamma 000} \cdots$$

se tiene

$$A_0 = p_0 p_{00} p_{000} \Pi_4 ,$$

$$A_1 < p_1 p_{10} p_{100} \Pi_4^{(1)} + p_0 p_{01} p_{010} \Pi_4^{(2)} + p_0 p_{00} p_{001} \Pi_4^{(3)}$$

$$A_2 < p_1 p_{11} p_{110} \Pi_4^{(4)} + p_1 p_{10} p_{101} \Pi_4^{(5)} + p_0 p_{01} p_{011} \Pi_4^{(6)}$$

$$A_3 < p_1 p_{11} p_{111} \Pi_4^{(7)} .$$

Sea P_4 el más pequeño de los productos Π_4. Las desigualdades anterio- dan

$$A_0 + A_1 + A_2 + A_3 > P_4(p_0 p_{00} p_{000} + \cdots + p_1 p_{11} p_{111})$$

El paréntesis contiene los productos que corresponden a todas las posibles combinaciones de índices, por lo tanto, es igual a

$$p_3 + q_3 = 1$$

Así que

$$S_4 = A_0 + A_1 + A_2 + A_3 > P_4$$

Esta desigualdad es evidentemente general. Se tiene de la misma forma

$$S_k > P_k \tag{1.10}$$

donde P_k designa el más pequeño de los productos

$$\Pi_k = p_{\alpha_1\alpha_2 \cdots \alpha_{k-1} 0} p_{\alpha_1\alpha_2 \cdots \alpha_{k-1} 00} p_{\alpha_1\alpha_2 \cdots \alpha_{k-1} 000} \cdots$$

Tenemos por hipótesis que

$$p_{\alpha_1\alpha_2 \cdots \alpha_{n-1} 1} < p''_n$$

y

$$p_{\alpha_1\alpha_2 \cdots \alpha_{n-1}0} + p_{\alpha_1\alpha_2 \cdots \alpha_{n-1}1} = 1$$

por lo tanto

$$p_{\alpha_1\alpha_2 \cdots \alpha_{n-1}0} > 1-p''_n$$

que conduce a la desigualdad

$$\Pi_k > \prod_{n=k}^{\infty} (1-p''_n)$$

Si la serie $\sum_n p''_n$ es convergente, el producto

$$\prod_{n=k}^{\infty} (1-p''_n)$$

también converge a un número distinto de cero, (véase lema 1 en el apéndice) por lo que podemos escoger el número k de manera que este producto sea superior a $1-\varepsilon$, cualquiera que sea ε. De aquí que

$$S_k > 1-\varepsilon$$

que es lo que se quería demostrar.

Para el caso de divergencia, observemos que, si la serie $\sum_n p'_n$ diverge, todos los productos infinitos que figuran dentro de $A_0, A_1, \ldots,$ son nulos. Estos productos son de la forma

$$P = p_\alpha p_{\alpha\beta} \cdots p_{\alpha\beta \cdots \lambda} \cdots$$

y se cumple que

$$p_{\alpha_1\alpha_2 \cdots \alpha_{n-1}1} > p'_n$$

de donde

$$p_{\alpha_1\alpha_2 \cdots \alpha_{n-1}0} < 1-p'_n$$

Para cualquier Π de la forma

$$\Pi = p_{\alpha\beta\gamma \cdots \lambda 0} p_{\alpha\beta\gamma \cdots \lambda 00} p_{\alpha\beta\gamma \cdots \lambda 000} \cdots$$

se tendrá, según lo anterior

$$\Pi < \prod_{n=k}^{\infty} (1-p'_n)$$

Si la serie $\sum_n p'_n$ es convergente, el producto infinito

$$\prod_{n=k}^{\infty} (1-p'_n)$$

tiene un valor nulo, esto da

$$\Pi = 0 \quad y \quad P = 0$$

1.4 Análisis del caso dependiente.

Esta nueva prueba de las probabilidades numerables, que Borel da, para el caso dependiente; usa de manera esencial la ley general de "probabilidades compuestas" (en términos de probabilidades condicionales) cuya expresión analítica es la siguiente: si denotamos por $E_n=1$ si se obtuvo el caso favorable, $E_n=0$ si se obtuvo caso desfavorable en el n-ésimo ensayo y denotamos por x a la sucesión infinita de ensayos; entonces, para cualquier sucesión infinita de ensayos x, la probabilidad que tenga valores iniciales fijos $E_i=m_i$, donde $m_i= 0$ o 1 para i = 1, 2,..., n, está dada por

$$P\left[E_1=m_1,E_2=m_2,\ldots,E_n=m_n\right] = \prod_{k=1}^{n} P\left[E_k=m_k \mid E_1=m_1,\ldots,E_{k-1}=m_{k-1}\right]$$

La extensión numerable que Borel da a esta ley en 1912 se puede enunciar de la siguiente manera. Sea x una sucesión infinita de ensayos, la probabilidad que x tenga los valores $E_{n+1}=m_{n+1}$, $E_{n+2}=m_{n+2},\ldots$, dado que sus primeros valores fueron $E_1=m_1$, $E_2=m_2,\ldots,E_n=m_n$ está dada por el producto infinito

$$\prod_{k=1}^{\infty} P\left[E_{n+k}=m_{n+k} \mid E_1=m_1,\ldots,E_{n+k-1}=m_{n+k-1}\right]$$

Borel muestra, aplicando directamente esta ley generalizada de probabilidades compuestas, que si $\sum_n p''_n$ converge entonces dado $\varepsilon>0$ existe k de tal manera que

$$A_0 + A_1 + \cdots + A_{k-1} > 1-\varepsilon$$

se cumple. La suma $A_0 + A_1 + \cdots + A_{k-1}$ es el valor de la probabilidad de obtener a lo más k-1 casos favorables.

La desigualdad (1.10) se puede escribir como

$$(1.11) \qquad A_0+A_1+\cdots+A_{k-1} > \prod_{i=k}^{\infty} P\left[E_i=0 \mid E_1=m_1,\ldots,E_{k-1}=m_{k-1},E_k=0,\ldots,E_{i-1}=0\right]$$

donde escogemos el mínimo de estos productos infinitos de los 2^{k-1} que se pueden obtener variando $m_j = 0$ o 1 para $j = 1,\ldots, k-1$. Por hipótesis se cumple

$$(1.12) \qquad p_n' < P\left[E_n=1 \mid E_1=m_1,E_2=m_2,\ldots,E_{n-1}=m_{n-1}\right] < p_n''$$

para cualesquiera valores de $m_j = 0$ o 1 $j = 1, 2,\ldots, n-1$ de (1.12) se sigue

$$1-p_n'' < 1-P\left[E_n=1 \mid E_1=m_1,\ldots,E_{n-1}=m_{n-1}\right]$$

$$(1.13) \qquad 1-p_n'' < P\left[E_n=0 \mid E_1=m_1,\ldots,E_{n-1}=m_{n-1}\right]$$

De (1.11) y (1.13) obtenemos

$$A_0+A_1+\cdots+A_{k-1} > \prod_{i=k}^{\infty} (1-p_i'')$$

como $\sum_i p_i'' < \infty$, entonces se puede escoger k de tal manera que este producto infinito sea mayor que $1-\varepsilon$ como se deseaba, demostrando que

$$A_0+A_1+\cdots+A_{k-1}+\cdots = 1.$$

En forma similar, si $\sum_i p_i'$ diverge, de (1.12) se sigue que

$$P\left[E_n=0|E_1=m_1,\ldots,E_{n-1}=m_{n-1}\right] < 1-p_n'$$

de modo que

$$P\left[E_n=0,E_{n+1}=0,\ldots,E_M=0|E_1=m_1,\ldots,E_{n-1}=m_{n-1}\right] < (1-p_n') \cdots (1-p_M')$$

Si tomamos el límite cuando $M \to \infty$, nos queda

$$P\left[E_n=0,E_{n+1}=0,\ldots|E_1=m_1,\ldots,E_{n-1}=m_{n-1}\right] = 0$$

Así, todos los productos que figuran dentro de las expresiones para $A_0, A_1,\ldots, A_{k-1}$ son nulos y por lo tanto también estos últimos valores, obteniendose así

$$A_0 + A_1 + \cdots + A_{k-1} = 0$$

para toda k.

En resumen, Borel obtiene:

$$(1.14)\quad \begin{cases} \text{Si } \sum\limits_n p_n'' \text{ converge, entonces } A_0 + A_1 + \cdots + A_{k-1} + \cdots = 1. \\ \\ \text{Si } \sum\limits_n p_n' \text{ diverge, entonces } A_0 + A_1 + \cdots + A_{k-1} + \cdots = 0. \end{cases}$$

o equivalentemente,

$$\text{Si } \sum_n p_n'' \text{ converge, entonces } A_\infty = 0.$$

Si $\sum_n p'_n$ diverge, entonces $A_\infty = 1$.

En conclusión, Borel en 1912 presenta una prueba de sus probabilidades numerables más general, basándose en un nuevo método más poderoso. (probabilidad compuesta numerable.)

1.5 Lema de Borel - Cantelli.

El resultado (1.14) fué conocido como la Ley Cero - Uno de Borel, este resultado ha sufrido algunas generalizaciones a partir de su aparición en 1909. En particular cuando $\sum_n p_n$ converge, entonces $A_\infty = 0$; se ha establecido sin necesidad de recurrir a la independencia entre ensayos (atribuido a Cantelli) y los ensayos de la exposición original de Borel han sido cambiados por eventos en un espacio de probabilidad. Se conoce modernamente a la Ley Cero - Uno de Borel como Lema de Borel - Cantelli.

Lema de Borel - Cantelli. Sea $(\Omega, \mathcal{B}, P)$ un espacio de probabilidad, de modo que $\mathcal{B}$ es un σ-campo de conjuntos de Ω y P es una medida aditiva numerable, no negativa y normalizada, definida en los conjuntos de $\mathcal{B}$. Sean $A_1, A_2, \ldots, A_n, \ldots$ una colección de eventos en $\mathcal{B}$.

(a) Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, entonces $P(\text{Lim sup}_n A_n) = 0$.

(b) Si los eventos A_n son independientes, entonces

$$\sum_{n=1}^{\infty} P(A_n) = \begin{matrix}\text{finito}\\ \text{o}\\ +\infty\end{matrix} \quad \text{sí y sólo sí} \quad P(\text{Lim sup}_n A_n) = \begin{matrix}0\\ \text{o}\\ 1\end{matrix}$$

Donde $\text{Lim sup}_n A_n = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n$ y $\text{Lim inf}_n = \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_n$

Prueba: (a) Sea $A = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n$. Entonces podemos escribir $A = \bigcap_{N=1}^{\infty} B_N$, donde $B_N = \bigcup_{n=N}^{\infty} A_n$. Para todo entero positivo N se cumple, por subaditividad numerable,

$$P(B_N) \leq \sum_{n=N}^{\infty} P(A_n)$$

como $\sum_{n=1}^{\infty} P(A_n) < \infty$, el lado derecho puede hacerse arbitrariamente pequeño, escogiendo N suficientemente grande. De modo que

$$P(A) = \lim_{N \to \infty} P(B_N) = 0.$$

(b) Supongamos ahora que los eventos A_n son independientes. Es obvio de las definiciones que

$$(1.15) \qquad P(\text{Lim sup}_n A_n) = 1 - P(\text{Lim inf}_n A_n^c)$$

donde A_n^c denota el complemento de A_n.
Sin embargo

$$(1.16) \qquad P(\text{Lim inf}_n A_n^c) = \lim_{k \to \infty} P(\bigcap_{n=k}^{\infty} A_n^c) = \lim_{k \to \infty} \prod_{n=k}^{\infty} \left[1 - P(A_n)\right]$$

de (1.15), (1.16) y los lemas 1 y 2 del apéndice se concluye el resultado.

CAPITULO 2

Las Fracciones Continuadas

2.1 Generalidades.

Una fracción continuada es una expresión de la forma

$$(2.1) \qquad a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}$$

en donde, para nüestros propósitos, a_0, a_1, a_2,... son números naturales. Diremos que una fracción continuada es limitada si tiene una expresión de la forma

$$(2.2) \qquad a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}$$

y es ilimitada si tiene una expresión de la forma dada por (2.1). Podemos desarrollar una fracción ordinaria en fracción continuada limitada cuando se efectuan, sobre el numerador y el denominador de la fracción considerada, las operaciones de búsqueda del máximo común divisor. Sea por ejemplo la fracción

$$\frac{8645}{2823}$$

el máximo común divisor del numerador y del denominador está dado por las operaciones

	3	16	25	7
8645	2823	176	7	1
176	7	1	0	

Escribimos, por lo tanto

$$\frac{8645}{2823} = 3 + \cfrac{1}{16 + \cfrac{1}{25 + \cfrac{1}{7}}}$$

Para simplificar la notación conveniremos, desde ahora, en escribir la fracción continuada ilimitada (2.1) en la forma

(2.3) $$[a_0, a_1, a_2, \ldots]$$

y la fracción continuada limitada (2.2) en la forma

(2.4) $$[a_0, a_1, a_2, \ldots, a_n]$$

de modo que el número de términos que aparecen en la fracción continuada limitada es igual al número de símbolos entre paréntesis cuadrados.

Llamaremos residuo de orden k de la fracción continuada (2.4) a la fracción

$$\frac{P_k}{Q_k} = [a_0, a_1, a_2, \ldots, a_k] \qquad \text{con } (k \leq n)$$

De igual forma podemos referirnos a $\frac{P'_k}{Q_k}$, para cualquier $k \geq 0$, al residuo de orden k de la fracción continuada ilimitada (2.3). Cláramente cualquier residuo, de una fracción continuada limitada o ilimitada, es una fracción continuada.

Además, llamaremos a la fracción continuada

$$r_k = [a_k, a_{k+1}, \ldots, a_n]$$

el resto de la fracción continuada (2.4). Similarmente nos referiremos a la fracción continuada

$$r_k = [a_k, a_{k+1}, a_{k+2}, \ldots]$$

como el resto de la fracción continuada ilimitada (2.3).

Por definición se sigue que para una fracción continuada limitada

$$[a_0, a_1, \ldots, a_n] = [a_0, a_1, \ldots, a_{k-1}, r_k] \quad (0 \leq k \leq n)$$

y para una fracción continuada ilimitada

$$[a_0, a_1, a_2, \ldots] = [a_0, a_1, \ldots, a_{k-1}, r_k] \quad (k \geq 0).$$

Para $k = 0$ y $k = 1$ los residuos son

$$\frac{P_0}{Q_0} = a_0$$

$$\frac{P_1}{Q_1} = a_0 + \frac{1}{a_1} = \frac{a_0a_1 + 1}{a_1}$$

Tomemos

$$(2.5) \qquad \begin{cases} P_0 = a_0 & ; \quad Q_0 = 1 \\ P_1 = a_0a_1 + 1 & ; \quad Q_1 = a_1 \end{cases}$$

Así, para k = 2

$$\frac{P_2}{Q_2} = \frac{a_0(a_1 + \frac{1}{a_2}) + 1}{a_1 + \frac{1}{a_2}} = \frac{a_0(a_1a_2 + 1) + a_2}{a_1a_2 + 1} = \frac{(a_0a_1 + 1)a_2 + a_0}{a_1a_2 + 1}$$

de donde

$$P_2 = (a_0a_1 + 1)a_2 + a_0 = P_0 + a_2P_1$$

$$Q_2 = a_1a_2 + 1 \qquad = Q_0 + a_2Q_1 \quad .$$

Estas últimas ecuaciones de recurrencia, motivan el teorema siguiente:

Teorema 1. Para cualquier $k \geq 2$, las siguientes fórmulas de recurrencia se cumplen

$$(2.6) \qquad \begin{cases} P_k = P_{k-2} + a_kP_{k-1} \\ Q_k = Q_{k-2} + a_kQ_{k-1} \end{cases} \quad .$$

Prueba. Para $k = 2$ las fórmulas (2.6) se cumplen como hemos visto anteriormente. Supongamos que el resultado es válido para toda $k < n$, por lo tanto, se cumple que

$$\frac{P_{n-1}}{Q_{n-1}} = \frac{P_{n-3} + a_{n-1}P_{n-2}}{Q_{n-3} + a_{n-1}Q_{n-2}}$$

y en consecuencia

$$\frac{P_n}{Q_n} = \frac{P_{n-3} + (a_{n-1} + \frac{1}{a_n})\, P_{n-2}}{Q_{n-3} + (a_{n-1} + \frac{1}{a_n})\, Q_{n-2}}$$

$$= \frac{P_{n-2} + a_n\, (P_{n-3} + a_{n-1}P_{n-2})}{Q_{n-2} + a_n\, (Q_{n-3} + a_{n-1}Q_{n-2})}$$

$$= \frac{P_{n-2} + a_nP_{n-1}}{Q_{n-2} + a_nQ_{n-1}}$$

esto demuestra las fórmulas (2.6).

Teorema 2. Para toda $k \geq 1$ se cumple

$$P_{k-1}Q_k - P_k\, Q_{k-1} = (-1)^k \tag{2.7}$$

Prueba. Calculemos el determinante D_k siguiente

$$D_k = \begin{vmatrix} P_{k-1} & P_k \\ Q_{k-1} & Q_k \end{vmatrix} = \begin{vmatrix} P_{k-1} & P_{k-2} + a_kP_{k-1} \\ Q_{k-1} & Q_{k-2} + a_kQ_{k-1} \end{vmatrix}$$

$$= \begin{vmatrix} P_{k-1} & P_{k-2} \\ & \\ Q_{k-1} & Q_{k-2} \end{vmatrix} = -D_{k-1}$$

y de (2.5) se obtiene

$$D_1 = \begin{vmatrix} a_0 & a_0a_1 + 1 \\ & \\ 1 & a_1 \end{vmatrix} = -1$$

combinando estos dos resultados, resulta que

$$D_k = P_{k-1}Q_k - P_kQ_{k-1} = (-1)^k.$$

Como consecuencia inmediata de este resultado, se sigue el siguiente corolario.

Corolario 1. Todo residuo de orden k de una fracción continuada es irreducible, es decir, P_k y Q_k son primos relativos.

Prueba. Cualquier divisor común de P_k y Q_k también lo es de $Q_kP_{k-1} - P_kQ_{k-1}$, pero de (2.7), este divisor es la unidad.

Por lo tanto, P_k y Q_k son primos relativos.

Teorema 3. Para cualquier sucesión a_0, a_1, a_2,... de números naturales, la expresión

$$\frac{P_k}{Q_k} = [\, a_0, a_1, a_2, \ldots, a_k \,]$$

tiene un límite cuando $k \to \infty$.

Prueba. Dividiendo la ecuación (2.7) entre $-Q_k\, Q_{k-1}$ obtenemos,

$$\frac{P_k}{Q_k} - \frac{P_{k-1}}{Q_{k-1}} = \frac{(-1)^{k-1}}{Q_k\, Q_{k-1}}$$

de la misma manera

$$\frac{P_{k-1}}{Q_{k-1}} - \frac{P_{k-2}}{Q_{k-2}} = \frac{(-1)^{k-2}}{Q_{k-1}\, Q_{k-2}}$$

hasta llegar por último a

$$\frac{P_1}{Q_1} - \frac{P_0}{Q_0} = \frac{1}{Q_1\, Q_0}$$

sumando miembro a miembro todas estas ecuaciones y pasando al final $\frac{P_0}{Q_0}$ al segundo miembro , se tiene

$$\frac{P_k}{Q_k} = \frac{P_0}{Q_0} + \frac{1}{Q_0\, Q_1} - \frac{1}{Q_1\, Q_2} + \frac{1}{Q_2\, Q_3} - \ldots + \frac{(-1)^{k-1}}{Q_{k-1}\, Q_k} \tag{2.8}$$

de manera que el k- ésimo residuo es la suma de los primeros k términos de una serie alternante de término general

$$\frac{(-1)^{n-1}}{Q_{n-1}\, Q_n}$$

para demostrar que esta serie es convergente, tenemos que :

$$Q_k = Q_{k-2} + a_k Q_{k-1}$$

$$\geq Q_{k-2} + Q_{k-1}$$

así

$$Q_{k-1} > Q_{k-2}$$

por lo tanto

$$Q_k > 2\, Q_{k-2}$$

y finalmente

$$Q_k\, Q_{k-1} > 2\, Q_{k-1}\, Q_{k-2} \, ,$$

por consiguiente , el valor absoluto de un término cualquiera es inferior a la mitad del término precedente.Esta serie es por lo tanto, convergente y su suma es el límite de $\frac{P_k}{Q_k}$ cuando $k \to \infty$.

La relación (2.8) muestra que el valor de la fracción continuada ilimitada está comprendida entre dos residuos sucesivos, los residuos de orden par son inferiores a la fracción continuada ilimitada y los residuos de orden impar son superiores.

Teorema 4. A todo número real positivo x le corresponde una fracción continuada única cuyo valor es este número . Esta fracción continuada es limitada si x es racional e ilimitada si x es irracional.

Prueba. Denotemos por a_0 el mayor entero menor o igual a x . Si x no es entero , entonces se puede escribir

$$x = a_0 + \frac{1}{r_1}$$

claramente $r_1 > 1$ porque

$$\frac{1}{r_1} = x - a_0 < 1$$

En general denotaremos por a_n el entero mayor , menor o igual a r_n (si r_n no es entero) ; el número r_{n-1} está determinado por la relación

$$r_n = a_n + \frac{1}{r_{n-1}}$$

Este proceso puede continuar hasta algún número r_m entero con $r_m > 1$ ($m \geq 1$)

Si x es racional , entonces claramente lo son todos los r_n . Es fácil ver que en este caso el proceso terminará después de un número finito de pasos . Si por ejemplo , $r_n = \frac{a}{b}$ entonces

$$r_n - a_n = \frac{a - ba_n}{b} = \frac{c}{b} < 1$$

$$r_n - a_n = \frac{1}{r_{n+1}}$$

con $c < b$

y por lo tanto

$$r_{n+1} = \frac{b}{c}$$

(suponiendo que c $\neq$ 0, es decir que r_n no es un entero) .

De aquí , r_{n+1} tiene un denominador menor que r_n ; así después de un numero finito de pasos deberá llegarse a un entero $r_n = a_n$. En este caso el número x estará representado por una fracción continuada cuyo último elemento es $a_n = r_n$ con $r_n > 1$.

Si el número x es irracional , entonces los r_n también lo son y nuestro proceso continuará indefinidamente. Escribiendo

$$\frac{P_n}{Q_n} = [a_0, a_1, a_2, \ . \ . \ . \ . \ , a_n]$$

sabemos que

$$x = \frac{P_{n-2} + r_n P_{n-1}}{Q_{n-2} + r_n Q_{n-1}}$$

y

$$\frac{P_n}{Q_n} = \frac{P_{n-2} + a_n P_{n-1}}{Q_{n-2} + a_n Q_{n-1}}$$

así, tomando la diferencia entre estos dos números ,

$$x - \frac{P_n}{Q_n} = \frac{(P_{n-1} Q_{n-2} - P_{n-2} Q_{n-1}) (r_n - a_n)}{(r_n Q_{n-1} + Q_{n-2}) (a_n Q_{n-1} + Q_{n-2})}$$

y sacando valor absoluto obtenemos ,

$$| x - \frac{P_n}{Q_n} | < \frac{1}{(r_n Q_{n-1} + Q_{n-2}) (a_n Q_{n-1} + Q_{n-2})}$$

$$< \frac{1}{Q_n^2}$$

Por lo tanto ,

$$\text{cuando } n \to \infty \qquad \frac{P_n}{Q_n} \to x$$

Claramente, esto significa que el valor de la fracción continuada infinita $[a_0, a_1, a_2, \ldots]$ es x .

Hemos demostrado hasta aquí que el número x puede siempre representarse por una fracción continuada , ésta termina si x es racional y es infinita si x es irracional . Sólamente resta mostrar que el desarrollo obtenido es único.

Sea $\qquad x = [a_0, a_1, a_2, \ldots] = [a_0', a_1', a_2', \ldots]$,

donde estas fracciones continuadas pueden ser limitadas o ilimitadas . Denotemos por $[\alpha]$ el mayor entero que no excede al número α . Es claro que $a_0 = [x]$ y $a_0' = [x]$, por lo cual $a_0 = a_0'$. Supongamos válido que $a_i' = a_i$ para toda $i = 0, 1, \ldots, n$, entonces tenemos que

$$\left.\begin{aligned} P_i' &= P_i \\ Q_i' &= Q_i \end{aligned}\right\} \quad \text{para } i = 0, \ldots, n$$

y

$$x = \frac{r_{n+1} P_n + P_{n-1}}{r_{n+1} Q_n + Q_{n-1}} = \frac{r'_{n+1} P'_n + P'_{n-1}}{r'_{n+1} Q'_n + Q'_{n-1}} = \frac{r'_{n+1} P_n + P_{n-1}}{r'_{n+1} Q_n + Q_{n-1}}$$

se sigue que $r_{n+1} = r'_{n+1}$ y como $a_{n+1} = [\ r_{n+1}\ ]$ y $a'_{n+1} = [\ r'_{n+1}\ ]$ entonces $a'_{n+1} = a_{n+1}$ es decir , las dos fracciones continuadas son idénticas.

2.2 Aplicación de Borel de las probabilidades numerables a las fracciones continuadas.

Consideremos la fracción continuada ilimitada

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_n + \cdots}}}$$

Para simplificar la notación escribiremos

$$x = [\, a_1, a_2, \ldots \,]$$

Nos proponemos buscar la probabilidad $P_{i,k}$ para que a_i sea igual al número k . Aquí $P_{i,k}$ es la medida del conjunto de los números comprendidos entre 0 y 1 para los cuales a_i es igual a k^*.

Se cumple evidentemente, puesto que a_i varía desde uno hasta infinito que,

$$\sum_{k=1}^{\infty} p_{i,k} = 1.$$

Cálculemos $p_{i,k}$ para los primeros valores de i. Para que a_i sea igual a la unidad, es necesario y suficiente que x cumpla con

$$\frac{1}{2} < x \leq 1$$

y así, x deberá estar comprendido dentro del intervalo $(\frac{1}{2}, 1]$. La probabilidad buscada es igual a la longitud de este intervalo

$$p_{1,1} = 1 - \frac{1}{2} = \frac{1}{2} .$$

De igual manera, a_1 es igual a 2 si x verifica la doble desigualdad

$$\frac{1}{3} < x \leq \frac{1}{2}.$$

Por lo tanto

$$p_{1,2} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

---

* Ya que $p_{i,k}$ es la unión numerable de intervalos ajenos; aquí Borel, está pensando la probabilidad $p_{i,k}$ como una medida, esto es, si un conjunto de puntos en el intervalo $[0,1]$ está formado por los puntos de una infinidad numerable de intervalos sin partes comunes, su medida es por definición, la suma de las longitudes de estos intervalos

En forma general, a_1 es igual a k si cumple

$$\frac{1}{k+1} < x \leq \frac{1}{k}$$

y por lo tanto,

$$p_{1,k} = \frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}$$

además se verifica

$$\sum_{k=1}^{\infty} p_{1,k} = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right) = 1.$$

Calculemos ahora las probabilidades relativas a a_2. Supongamos primero que a_1 es igual a n, si a_2 es igual a k, x estará comprendida entre las fracciones*

$$[n,k] \quad \text{y} \quad [n,k+1]$$

es decir, dentro del intervalo de longitud

$$\lambda_n = \frac{1}{n + \frac{1}{k+1}} - \frac{1}{n + \frac{1}{k}} = \frac{1}{[(k+1)n + 1]\,[kn + 1]}$$

---

* Obsérvese que aquí se está pidiendo que se cumpla simultáneamente que $a_1 = n$ y $a_2 = k$. Para ver que las fracciones continuadas que cumplen con ambas condiciones están entre $[n,k]$ y $[n,k+1]$, obsérvese la demostración del teorema 3.

λ_n representa la probabilidad que a_1 sea igual a n y a_2 a k. Es suficiente, por lo tanto, hacer la suma de los números λ_n para todos los valores de $n = 1, 2, \ldots$, para obtener el valor de $p_{2,k}$.

Tenemos que

$$p_{2,k} = \frac{1}{k(k+1)} \sum_{n=1}^{\infty} \frac{1}{(n + \frac{1}{k})(n + \frac{1}{k+1})} = \frac{A_k}{k(k+1)}$$

tomando $$A_k = \sum_{n=1}^{\infty} \frac{1}{(n + \frac{1}{k})(n + \frac{1}{k+1})}$$

El número A_k es inferior a la suma de la serie $\sum_{n=1}^{\infty} \frac{1}{n^2}$, suma que es igual a $\frac{\pi^2}{6}$, y superior a

$$\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+\frac{1}{2})} > \sum_{n=1}^{\infty} \frac{1}{(n + 1)^2} = \frac{\pi^2}{6} - 1$$

Por consiguiente, A_k está comprendida entre $\frac{\pi^2}{6}$ y $\frac{\pi^2}{6} - 1$ y tiende hacia $\frac{\pi^2}{6}$ cuando k aumenta indefinidamente. A consecuencia, para valores grandes de k, la probabilidad $p_{2,k}$ es mayor a la probabilidad $p_{1,k}$; por el contrario, se tiene, para valores pequeños de k,

$$p_{2,k} < p_{1,k}$$

Busquemos ahora, la probabilidad para que se tenga a_3 igual a k si suponemos a_1 igual a n y a_2 igual a n'. Para que esto suceda x debe estar comprendida entre las fracciones continuadas

$$[\, n, n', k+1 \,] \quad y \quad [\, n, n', k \,]$$

o

$$\frac{n'(k+1)+1}{(nn'+1)(k+1)+n} < x \leq \frac{n'k+1}{(nn'+1)k+n}$$

por lo tanto, x está dentro de un intervalo de longitud

$$\frac{1}{[(nn'+1)(k+1)+n\,]\;[(nn'+1)k+n\,]}$$

y la suma infinita sobre todos los valores posibles de n y n' nos da el valor buscado

$$p_{3,k} = \frac{}{k(k+1)} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1}{(nn'+1+\frac{n}{k+1})(nn'+1+\frac{n}{k})}$$

esta serie doble, comparable a la serie

$$\sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \frac{1}{n^2 n'^2}$$

es convergente.

El método hasta aquí empleado, nos da un medio teóricamente simple pero poco cómodo, para calcular las $p_{i,k}$ sucesivas, sin embargo, usando algunos de los resultados de la sección anterior podemos proseguir con nuestro estudio.

La condición necesaria y suficiente para que a_n sea igual a k, es que x esté comprendida entre las fracciones $[a_1, a_2, \ldots, a_{n-1}, k]$ y $[a_1, a_2, \ldots, a_{n-1}, k+1]$ o equivalentemente entre

$$\frac{P_{n-3} + (a_{n-1} + \frac{1}{k})P_{n-2}}{Q_{n-3} + (a_{n-1} + \frac{1}{k})Q_{n-2}} \quad y \quad \frac{P_{n-3} + (a_{n-1} + \frac{1}{k+1})P_{n-2}}{Q_{n-3} + (a_{n-1} + \frac{1}{k+1})Q_{n-2}}$$

que podemos poner como

$$\frac{P_{n-2} + kP_{n-1}}{Q_{n-2} + kQ_{n-1}} \quad y \quad \frac{P_{n-2} + (k+1)P_{n-1}}{Q_{n-2} + (k+1)Q_{n-1}}$$

este intervalo tiene una longitud

$$l_k = \left| \frac{P_{n-2} + kP_{n-1}}{Q_{n-2} + kQ_{n-1}} - \frac{P_{n-2} + (k+1)P_{n-1}}{Q_{n-2} + (k+1)Q_{n-1}} \right|$$

teniendo en cuenta la relación (2.7) se sigue

$$\left| P_{n-2}Q_{n-1} - P_{n-1}Q_{n-2} \right| = 1$$

y

$$l_k = \frac{1}{(Q_{n-2} + kQ_{n-1})(Q_{n-2} + (k+1)Q_{n-1})}$$

o

$$l_k = \frac{1}{Q^2_{n-1}(k + \frac{Q_{n-2}}{Q_{h-1}})(k + 1 + \frac{Q_{n-2}}{Q_{n-1}})}$$

La suma de las longitudes de todos los intervalos que se obtienen para todos los valores posibles de a_1, a_2,..., a_{n-1} es la probabilidad $p_{n,k}$. De la misma manera $p_{n,k+1}$ es la suma de todos los intervalos de longitud

$$l_{k+1} = \frac{1}{Q^2_{n-1}(k + 1 + \frac{Q_{n-2}}{Q_{n-1}})(k + 2 + \frac{Q_{n-2}}{Q_{n-1}})}$$

La razón de las longitudes de dos intervalos, correspondientes a un mismo sistema de valores de $a_1, a_2, \ldots, a_{n-1}$, es

$$\frac{l_{k+1}}{l_k} = \frac{k + \frac{Q_{n-2}}{Q_{n-1}}}{k + 2 + \frac{Q_{n-2}}{Q_{n-1}}}$$

Para un valor k dado, esta razón satisface la doble desigualdad

$$\text{(2.9)} \qquad \frac{k}{k+2} < \frac{l_{k+1}}{l_k} < \frac{k+1}{k+3}$$

ya que la razón $\frac{Q_{n-2}}{Q_{n-1}}$ varía entre 0 y 1, independientemente de los valores posibles de $a_1, a_2, \ldots, a_{n-1}$ (esto se sigue de la fórmula de recurrencia (2.6)).

Multiplicando (2.9) por l_k y sumando para todos los valores posibles de $a_1, a_2, \ldots, a_{n-1}$ obtenemos

$$\frac{k}{k+2}\, p_{n,k} < p_{n,k+1} < \frac{k+1}{k+3}\, p_{n,k}$$

y equivalentemente

$$\frac{k}{k+2} < \frac{p_{n,k+1}}{p_{n,k}} < \frac{k+1}{k+3}$$

Se obtiene por lo tanto, por recurrencia

$$\frac{(k-1)(k-2)\cdots 2\cdot 1}{(k+1)k\cdots 4\cdot 3} < \frac{p_{n,k}}{p_{n,1}} < \frac{k(k-1)\cdots 3\cdot 2}{(k+2)(k+1)\cdots 5\cdot 4}$$

es decir

(2.10) $$\frac{2}{k(k+1)} < \frac{p_{n,k}}{p_{n,1}} < \frac{6}{(k+1)(k+2)}$$

y utilizando la relación

(2.11) $$\sum_{k=1}^{\infty} p_{n,k} = 1$$

obtenemos

$$\sum_{k=1}^{\infty} \frac{2}{k(k+1)} < \frac{\sum_{k=1}^{\infty} p_{n,k}}{p_{n,1}} < \sum_{k=1}^{\infty} \frac{6}{(k+1)(k+2)}$$

es decir

$$2 < \frac{1}{p_{n,1}} < 3$$

finalmente

(2.12) $$\frac{1}{3} < p_{n,1} < \frac{1}{2}$$

Multiplicando miembro a miembro las desigualdades (2.10) y (2.12), obtenemos

(2.13) $$\frac{2}{3k(k+1)} < p_{n,k} < \frac{3}{(k+1)(k+2)} .$$

Si hacemos $k = 1$ en esta última doble desigualdad, obtenemos (2.12), así la desigualdad (2.13)-es general.

La probabilidad para que a_n sea superior a k es igual a

$$Q_{n,k} = p_{n,k+1} + p_{n,k+2} + \cdots$$

De (2.13), $Q_{n,k}$ satisface la doble desigualdad

$$\frac{2}{3}\sum_{\nu=k+1}^{\infty}\frac{1}{\nu(\nu+1)} < Q_{n,k} < 3\sum_{\nu=k+2}^{\infty}\frac{1}{\nu(\nu+1)}$$

o

$$\frac{2}{3(k+1)} < Q_{n,k} < \frac{3}{k+2}$$

Sea $\phi(n)$ una función creciente de n, tal que la serie

$$\sum_{n=1}^{\infty}\frac{1}{\phi(n)} \tag{2.14}$$

sea convergente. La serie

$$\sum_{n=1}^{\infty} Q_{n,\phi(n)}$$

es también convergente. Se concluye la proposición siguiente:

Proposición 1. (Teorema de las fracciones continuadas) Sea $\phi(n)$ una función tal que la serie (2.14) sea convergente, la probabilidad para que se tenga, para una infinidad de valores de n

$$a_n > \phi(n)$$

es igual a cero. Dicho de otra manera, existe una probabilidad igual a uno para que se tenga, a partir de un valor finito de n:

$$a_n < \phi(n).$$

Cualquiera que sea la función $\phi(n)$ tal que la serie (2.14) sea convergente, existe una función $\psi(n)$ que satisface la misma condición y cumple eon

$$\lim_{n\to\infty} \frac{\psi(n)}{\phi(n)} = 0.$$

de dónde se concluye que, si la serie (2.14) es convergente, existe una probabilidad igual a uno para que se tenga

$$\lim_{n\to\infty} \frac{a_n}{\phi(n)} = 0.$$

Si la serie (2.14) es divergente, no se puede afirmar que el límite anterior es igual a infinito, pero se cumplirá la relación

$$\overline{\text{Lim}} \frac{a_n}{\phi(n)} = \infty$$

donde $\overline{\text{Lim}}$ denota al límite superior de una sucesión.

2.3 Análisis de la aplicación de Borel a las fracciones continuadas.

La sección (2.2) es la exposición que Borel da de su capítulo III en el artículo de 1909. Una observación importante que podemos hacer de esta exposición es la siguiente: El conjunto de números $x \in [0,1]$ que satisfacen $a_n(x) = k$ es una unión de intervalos y no es por sí mismo un intervalo. La longitud de cada uno de estos intervalos (que Borel llama l_k) se calcula en términos de $P_n = P_n(a_1, a_2, \dots, a_n)$ y $Q_n = Q_n(a_1, a_2, \dots, a_n)$, esto lo podemos ver poniendo la doble desigualdad (2.9) en la forma

$$\frac{k}{k+2} < \frac{P\left[a_1=m_1,\ a_2=m_2,\ldots,\ a_{n-1}=m_{n-1},\ a_n=k+1\right]}{P\left[a_1=m_1, a_2=m_2,\ldots, a_{n-1}=m_{n-1}, a_n=k\ \right]} < \frac{k+1}{k+3}$$

donde los m_i son enteros positivos fijos, si multiplicamos esta desigualdad por $P\left[a_1=m_1,\ a_2=m_2,\ldots,\ a_{n-1}=m_{n-1},\ a_n=k\ \right]$ y sumamos sobre todos los valores posibles de $m_1,\ m_2,\ldots,m_{n-1}$ obtenemos

$$\frac{k}{k+2}\, p_{n,k} < p_{n,k+1} < \frac{k+1}{k+3}\, p_{n,k}$$

Esto se sigue, puesto que,

$$\sum_{m_1=1}^{\infty}\ \sum_{m_2=1}^{\infty} \cdots \sum_{m_{n-1}=1}^{\infty} P\left[a_1=m_1,\ldots,\ a_{n-1}=m_{n-1},\ a_n=k\ \right]= P\left[a_n=k\ \right]$$

ya que en este caso Borel si está pensando a la probabilidad como una medida, por lo tanto, la aditividad numerable puede ser usada sin ningún problema.

La exposición de Borel refleja una gran habilidad en la manipulación de las fracciones continuadas y de una manera bastante ingeniosa, construye una sucesión de ensayos y asocia caso favorable o desfavorable escogiendo una función entera fija $\phi(n)$ y definiendo para cada fracción continuada ilimitada x, el n-ésimo ensayo favorable o desfavorable si $a_n > \phi(n)$ o $a_n \leq \phi(n)$ respectivamente. Así, Borel ha asociado a cada fracción continuada ilimitada un ejemplo de una sucesión infinita de ensayos, donde cada ensayo únicamente tiene dos posibles resultados, entonces aplica su Ley Cero-Uno (sección 1.1) a esta colección de sucesiones de ensayos.

De su resultado sobre A_∞ Borel concluye:

$$P\ \{a_n(x) > \phi(n)\ \text{ una infinidad de veces }\} = 0 \text{ o } 1$$

dependiendo si $\Sigma \frac{1}{\phi(n)}$ converge o diverge. Adicionalmente, hace una mejor observación; si $\Sigma \frac{1}{\phi(n)}$ converge, entonces existe una función entera $\psi(n)$ tal que $\Sigma \frac{1}{\psi(n)}$ converge y se cumple con $\lim_{n\to\infty} \frac{\psi(n)}{\phi(n)} = 0$. Alternativamente, si $\Sigma \frac{1}{\phi(n)}$ diverge, existe una función $\psi(n)$ tal que $\Sigma \frac{1}{\psi(n)}$ diverge y se cumple con $\lim_{n\to\infty} \frac{\psi(n)}{\phi(n)} = \infty$.

Se sigue de la formulación original aplicada a $\psi(n)$ que el teorema puede ser formulado como:

$$P\,\{\overline{\lim}\, \frac{a_n}{\phi(n)} = 0\} = 1 \quad \text{si} \quad \Sigma \frac{1}{\phi(n)} \quad \text{converge}$$

o

$$P\,\{\overline{\lim}\, \frac{a_n}{\phi(n)} = \infty\} = 1 \quad \text{si} \quad \Sigma \frac{1}{\phi(n)} \quad \text{diverge.}$$

La aplicación de la ley Cero-Uno que Borel hace en las fracciones continuadas es errónea. La ley Cero-Uno fue obtenida (1909) considerando los ensayos independientes, pero en el caso de las fracciones continuadas los ensayos bajo consideración son dependientes, (esto se puede ver de manera clara tanto geométricamente como algebráicamente. Un ejemplo de esta dependencia fue dado por F. Bernstein y lo reproducimos en la siguiente sección.) la conclusión $A_\infty = 0$ puede justificarse con la generalización de Cantelli siempre que $\sum_n p_n$ sea convergente. Sin embargo, el resultado $A_\infty = 1$ si $\sum_n p_n$ diverge, requiere una generalización distinta de la discusión de Borel , original, del caso divergente. La generalización requerida del resultado $A_\infty = 1$ fue dado en 1911 por Felix Bernstein.

2.4 Crítica de Felix Bernstein a los resultados obtenidos por E. Borel acerca de las fracciones continuadas.

Felix Bernstein en su artículo de 1911 establece algunos resultados concernientes a las fracciones continuadas. A continuación hacemos una exposición de estos resultados.

Sea k un entero positivo tal que $k \geq 2$. Si para los valores fijos $a_1, a_2, \ldots, a_{n-1}$

$$a_n \geq k, \tag{2.15}$$

y por otro lado a_{n+p} para $p = 1, 2, \ldots$ variables, entonces las fracciones continuadas ilimitadas x están dentro del intervalo cuyos puntos extremos son las fracciones continuadas limitadas

$$[a_1, a_2, \ldots, a_{n-1}] \text{ y } [a_1, a_2, \ldots, a_{n-1}, k]$$

de longitud

$$l_{nk} = \left|\frac{P_{n-1}}{Q_{n-1}} - \frac{P_n}{Q_n}\right| = \frac{1}{Q_{n-1}Q_n}$$
$$= \frac{1}{Q_{n-1}(kQ_{n-1} + Q_{n-2})}$$

es decir

$$l_{nk} = \frac{1}{Q_{n-1}^2\left(k + \frac{Q_{n-2}}{Q_{n-1}}\right)} \tag{2.16}$$

de modo que

$$\frac{l_{nk}}{l_{n1}} = \frac{1 + \frac{Q_{n-2}}{Q_{n-1}}}{k + \frac{Q_{n-2}}{Q_{n-1}}}$$

de modo que

$$\frac{l_{nk}}{l_{n1}} = \frac{1 + \dfrac{Q_{n-2}}{Q_{n-1}}}{k + \dfrac{Q_{n-2}}{Q_{n-1}}}$$

y como

$$0 < \frac{Q_{n-2}}{Q_{n-1}} < 1$$

entonces

$$\min_{0 \leq x \leq 1} \frac{1 + x}{k + x} < \frac{l_{nk}}{l_{n1}} < \max_{0 \leq x \leq 1} \frac{1 + x}{k + x}$$

cuyos valores extremos se alcanzan en los puntos extremos que son $x = 0$ y $x = 1$. Es decir, se tiene

(2.17) $$\frac{1}{k} < \frac{l_{nk}}{l_{n1}} < \frac{2}{k+1}$$

Análogamente se tiene para $a_n < k$

(2.18) $$\frac{k-1}{k+1} < \frac{l_{n1} - l_{nk}}{l_{n1}} < \frac{k-1}{k}$$

Si multiplicamos las desigualdades (2.17) y (2.18) por l_{n1} y sumamos sobre todos los posibles valores de $a_1, a_2, \ldots, a_{n-1}$, obtenemos

(2.19) $$\frac{1}{k} \sum_i {}_i l_{n1} < \sum_i {}_i l_{nk} < \frac{2}{k+1} < \sum_i {}_i l_{n1}$$

Pero $\sum_i {}_i l_{n1}$ está tomada sobre todos los valores posibles de a_i para $i = 1, \ldots, n-1$ por lo que esta suma es igual a la longitud de todo el intervalo $[0,1]$ y $\sum_i {}_i l_{nk}$ es igual a $P[a_n \geq k]$ así que, obtenemos

(2.20) $$\frac{1}{k} < P\ [a_n \geq k] < \frac{2}{k+1}$$

en forma análoga podemos obtener la desigualdad

$$(2.21) \qquad \frac{k-1}{k+1} < P[a_n < k] < \frac{k-1}{k} .$$

Ahora en (2.19) cambiemos la primera n por $\bar{n} > n$ y pongamos $\bar{k}$ en lugar de k; tomemos luego la suma sobre todos los posibles valores de $a_1, \ldots, a_{\bar{n}-1}$ pero manteniendo siempre $a_n \geq k$, de modo que la suma $\Sigma\ _i 1_{\bar{n}1}$ es obviamente idéntica a $P[a_n \geq k]$ y obtenemos

$$(2.22) \qquad \frac{1}{\bar{k}} P[a_n \geq k] < P[a_n \geq k\ ,\ a_{\bar{n}} \geq \bar{k}] < \frac{2}{\bar{k}+1} P[a_n \geq k]$$

análogamente

$$(2.23) \qquad \frac{\bar{k}-1}{\bar{k}+1} P[a_n < k] < P[a_n < k\ ,\ a_{\bar{n}} < \bar{k}] < \frac{\bar{k}-1}{\bar{k}} P[a_n < k]$$

De (2.20), (2.21), (2.22) y (2.23) nos da

$$\frac{1}{\bar{k}}\,\frac{1}{k} < P[a_n \geq k\ ,\ a_{\bar{n}} \geq \bar{k}] < \frac{\bar{k}-1}{\bar{k}}\,\frac{k-1}{k} .$$

Este procedimiento puede ser empleado de manera general para mostrar que la probabilidad de que se cumplan simultáneamente las condiciones

$$a_{n_1} \geq k_{n_1},\ a_{n_2} \geq k_{n_2}, \ldots, a_{n_r} \geq k_{n_r}$$

satisface las desigualdades

$$(2.24) \qquad \frac{1}{k_{n_1}}\,\frac{1}{k_{n_2}} \cdots \frac{1}{k_{n_r}} < P[a_{n_1} \geq k_{n_1}, \ldots, a_{n_r} \geq k_{n_r}] < \frac{2}{(k_{n_1}+1)} \cdots \frac{2}{(k_{n_r}+1)}$$

Y similarmente

$$(2.25)\qquad (1 - \frac{2}{k_{n_1}+1}) \cdots (1 - \frac{2}{k_{n_r}+1}) < P[a_{n_1} < k_{n_1}, \ldots, a_{n_r} < k_{n_r}]$$

$$< (1 - \frac{1}{k_{n_1}}) \cdots (1 - \frac{1}{k_{n_r}}).$$

Luego, si $\phi(n)$ es una función mayor que cero y $\{a_{i_n}\}$ es una subsucesión de la sucesión de denominadores y ponemos

$$a_{i_n} < \phi(n)$$

en vez de la condición $a_{i_n} < k$, entonces de (2.25) tenemos

$$\prod_{\nu=n}^{n+r} \frac{\phi(\nu) - 1}{\phi(\nu) + 1} < P[a_{i_n} < \phi(n), \ldots, a_{i_{n+r}} < \phi(n+r)] < \prod_{\nu=n}^{n+r} (1 - \frac{1}{\phi(\nu)})$$

Si $\sum_{n=1}^{\infty} \frac{1}{\phi(n)}$ converge, entonces también $\sum_{n=1}^{\infty} \frac{2}{\phi(n)+1}$ es convergente y ambos productos $\prod_{\nu=n}^{\infty} (1 - \frac{1}{\phi(\nu)})$ y $\prod_{\nu=n}^{\infty} (1- \frac{2}{\phi(\nu)+1})$ son distintos de cero.

Si contrariamente $\sum_{n=1}^{\infty} \frac{1}{\phi(n)}$ diverge, entonces $\prod_{\nu=n}^{\infty} (1- \frac{1}{\phi(\nu)}) = 0$
Por lo tanto, concluimos con el siguiente teorema

Teorema 5. Una condición necesaria y suficiente para que se forme una probabilidad distinta de cero cuando $\phi(n)$ es cota superior o inferior para la subsucesión de denominadores $\{a_n\}$, tomada a partir de cierta n, radica en la convergencia de $\sum_{n=1}^{\infty} \frac{1}{\phi(n)}$. Si por el contrario, $\sum_{n=1}^{\infty} \frac{1}{(n)}$ diverge hay una probabilidad igual a 1 para que la desigualdad de acotamiento se rompa una infinidad de veces.

Después de enunciar el teorema anterior, Bernstein dice que el resultado $A_\infty = 0$ si $\sum_n \frac{1}{\phi(n)}$ converge, de Borel, está en contradicción con la primera parte del teorema 5.

" La razón de la contradicción tiene una importancia principal que queremos hacer notar explícitamente.

Consiste en el siguiente hecho:

Para nuestra probabilidad geométrica no vale la independencia de los casos particulares.

Para ello, basta construir un simple ejemplo. Denotemos por ${}_1p_1$ y ${}_2p_1$ a la probabilidad que se tenga $a_1 = 1$ y $a_2 = 2$ respectivamente y por p_{11}, p_{21} a la probabilidad que se tenga $a_1 = 1,\ a_2 = 1$ y $a_1 = 2,\ a_2 = 1$ respectivamente, de este modo si se aceptara la independencia de los casos particulares debería suceder que $p_{11}:p_{21} = {}_1p_1:{}_2p_1$

$$p_{11} = \frac{1}{1+\frac{1}{1+1}} - \frac{1}{1+\frac{1}{1+0}} = \frac{1}{6},$$

$$p_{21} = \frac{\quad}{2+\frac{1}{1+1}} - \frac{\quad}{2+\frac{1}{1+0}} = \frac{1}{15},$$

$${}_1p_1 = \frac{1}{1+0} - \frac{1}{1+1} = \frac{1}{2},$$

$${}_2p_1 = \frac{1}{2+0} - \frac{1}{2+1} = \frac{1}{6}.$$ " F. Bernstein (1911: 431)

Hay dos puntos esencialmente importantes en este trabajo de Bernstein.

El primero es su concepto de probabilidad geométrica; Bernstein llama probabilidad geométrica a las razones $\frac{l_{nk}}{l_{n1}}$, $\frac{l_{n1} - l_{nk}}{l_{n1}}$ de las desi--

gualdades (2.17) y (2.18). Si introducimos la notación

$$P[a_n \geq k|\ a_1=m_1,\ldots,a_{n-1}=m_{n-1}]$$

la cual denota, la probabilidad condicional de que $a_n \geq k$ dado que $a_i(x)=m_i$ para $i = 1,2,\ldots,n-1$, entonces (2.17) y (2.18) se pueden escribir como

$$(2.26) \qquad \frac{1}{k} < P[a_n \geq k|\ a_1=m_1,\ldots,a_{n-1}=m_{n-1}] < \frac{2}{k+1}$$

$$(2.27) \qquad \frac{k-1}{k+1} < P[a_n < k|\ a_1=m_1,\ldots,a_{n-1}=m_{n-1}] < \frac{k-1}{k}$$

respectivamente ya que

$$\frac{l_{nk}}{l_{n1}} = P[a_n \geq k|\ a_1=m_1,\ldots,a_{n-1}=m_{n-1}]$$

y

$$\frac{l_{n1} - l_{nk}}{l_{n1}} = P[a_n < k|\ a_1=m_1,\ldots,a_{n-1}=m_{n-1}]$$

de donde se obtienen las desigualdades globales (2.20) y (2.21) como promedios pesados de (2.26) y (2.27).

Las desigualdades (2.20) y (2.21) pueden compararse con las obtenidas por Borel

$$\frac{2}{3(k+1)} < P[a_n > k] < \frac{3}{k+2}$$

$$\frac{k-1}{k+2} < P[a_n \leq k] < \frac{3k+1}{3k+3}$$

La distinción entre éstas y las desigualdades de Bernstein (2.20) y

(2.21) es esencial porque Borel no ha calculado cotas para las probabilidades condicionales, solamente para probabilidades globales. Como es de esperarse, las cotas de Bernstein limitan mejor los valores reales de estas probabilidades.

El segundo punto importante en este trabajo,es el ingenioso argumento que Bernstein da para obtener las desigualdades (2.24) y (2.25). Parte de la desigualdad (2.19) tomando $\bar{n}$ por n y $\bar{k}$ por k. Escribiendo esta desigualdad en la notacion anterior, nos queda

$$\frac{1}{\bar{k}} P[a_{\bar{n}} \geq 1, a_{\bar{n}-1}=m_{\bar{n}-1},\ldots,a_1=m_1] < P[a_{\bar{n}} \geq \bar{k}, a_{\bar{n}-1}=m_{\bar{n}-1},\ldots,a_1=m_1]$$

$$< \frac{2}{\bar{k}+1} P[a_{\bar{n}} \geq 1, a_{\bar{n}-1}=m_{\bar{n}-1},\ldots,a_1=m_1]$$

se sigue que

$$\frac{1}{\bar{k}} \{ \sum_{m_1=1}^{\infty} \cdots \sum_{m_n=k}^{\infty} \cdots \sum_{m_{\bar{n}-1}}^{\infty} P[a_{\bar{n}} \geq 1;\ a_i=m_i,\ 1 \leq i \leq \bar{n}-1] \}$$

$$< \sum_{m_1=1}^{\infty} \cdots \sum_{m_n=k}^{\infty} \cdots \sum_{m_{\bar{n}-1}}^{\infty} P[a_{\bar{n}} \geq k; a_i=m_i,\ 1 \leq i \leq \bar{n}-1]$$

$$< \frac{2}{\bar{k}+1} \{ \sum_{m_1=1}^{\infty} \cdots \sum_{m_n=k}^{\infty} \cdots \sum_{m_{\bar{n}-1}}^{\infty} P[a_{\bar{n}} \geq 1; a_i=m_i, 1 \leq i \leq \bar{n}-1]\}$$

de donde obtenemos

$$\frac{1}{\bar{k}} P[a_n \geq k] < P[a_n \geq k\ ,\ a_{\bar{n}} \geq \bar{k}] < \frac{2}{\bar{k}+1} P[a_n \geq k\ ]$$

y usando (2.20) nos queda

$$\frac{1}{\bar{k}}\,\frac{1}{k} < P[a_n \geq k\ ,\ a_{\bar{n}} \geq \bar{k}] < \frac{2}{\bar{k}+1}\,\frac{2}{k+1}$$

De manera semejante se puede obtener

$$\frac{\bar{k}-1}{\bar{k}+1}\,\frac{k-1}{k+1} < P[a_n < k\ ,\ a_{\bar{n}} < \bar{k}] < \frac{\bar{k}-1}{\bar{k}}\,\frac{k-1}{k}$$

Una manera alternativa y más conveniente de usar este argumento de Bernstein para obtener en forma general las desigualdades (2.24) y (2.25) es la propuesta por J. Barone y A. Novikoff. De la ley de probabilidades compuestas

$$(2.28)\qquad P[a_1=m_1,\dots,a_n=m_n] = P[a_n=m_n \mid a_i=m_i, 1 \leq i \leq n-1]$$

$$\cdots P[a_2=m_2 \mid a_1=m_1]\, P[a_1=m_1]$$

denotemos el producto a la derecha por $\Pi_n (m_1,m_2,\dots,m_n)$. Fijemos $m_1,m_2,\dots,m_{n-1}$ y sumando ambos lados para todos los valores posibles de m_n variando desde un límite inferior fijo hasta infinito. Si el límite inferior es 1, ambos lados se simplifican y el resultado es la afirmación original con todos las referencias de a_n borradas, envolviendo Π_{n-1} en el lado derecho. Si el límite inferior es algún entero $k_n > 1$, el lado derecho queda

$$P[a_1=m_1,\dots,a_{n-1}=m_{n-1},a_n \geq k_n]$$

mientras que el lado derecho de (2.28) es igual al producto de Π_{n-1} por

$P[a_n \geq k_n \mid a_1=m_1,\dots,a_{n-1}=m_{n-1}]$ y utilizando (2.26) nos queda

$$\frac{1}{k_n}\,\underset{n-1}{\Pi} < P[a_i=m_i\ ,\ 1\leq\ i\ \leq n-1\ ,\ a_n \geq k_n] < \frac{2}{k_n+1}\,\underset{n-1}{\Pi}$$

En forma similar estimamos $\underset{n-1}{\Pi}$ sumando sobre m_{n-1} de algún límite inferior k_{n-1} a infinito, quitando la referencia a a_{n-1} si $k_{n-1}=1$ o también obteniendo

$$\frac{1}{k_{n-1}}\,\underset{n-2}{\Pi} < \underset{n-1}{\Pi} < \frac{2}{k_{n-1}+1}\,\underset{n-2}{\Pi}$$

si el límite inferior $k_{n-1} > 1$. Procediendo de esta manera y suponiendo que $a_{n_1}, a_{n_2},\dots,a_{n_r}$ tienen límites inferiores $k_{n_1}, k_{n_2},\dots,k_{n_r}$ llegamos a la desigualdad (2.24), donde los índices intermedios ausentes corresponden a sumas con límite inferior igual a 1. De manera análoga podemos obtener la desigualdad (2.25).

Para obtener el resultado de fracciones continuadas de Borel, es suficiente sustituir en (2.25) $\phi(n)$ por k_n. Si $\sum\limits_{n=1}^{\infty}\frac{1}{\phi(n)}$ diverge,

$$P[a_\nu < \phi(\nu),\ \nu=n,n+1,\dots] = \lim_{r\to\infty} P[a_\nu < \phi(\nu),\ \nu=n,n+1,\dots,r] = 0.$$

que es precisamente el caso A_∞ de Borel.

Si $\sum\limits_{n=1}^{\infty}\frac{1}{\phi(n)}$ converge, entonces (2.25) afirma que

$$P[a_\nu < \phi(\nu),\ \nu=n,n+1,\dots] = \lim_{r\to\infty}\ P[a_\nu < \phi(\nu),\ \nu=n,n+1,\dots,r]$$

es positivo (menor que 1) y está acotado entre los valores

$$\prod_{\nu=n}^{\infty}\left(1-\frac{2}{\phi(\nu)+1}\right) \quad y \quad \prod_{\nu=n}^{\infty}\left(1-\frac{1}{\phi(\nu)}\right)$$

En este caso para obtener el resultado de Borel, es necesario calcular la probabilidad que $a_\nu < \phi(\nu)$ se cumpla, esto es, la probabilidad de la unión de los conjuntos $\{a_\nu < \phi(\nu),\ \nu \geq n\}$ sobre todos los valores enteros de ν, esta probabilidad es uno ya que para valores muy grandes de n los productos infinitos anteriores tienden a 1.

2.5 Respuesta de Borel a la crítica de Bernstein.

Después de que Bernstein critica los resultados de Borel (mostrando que sus resultados de probabilidades numerables no podían ser aplicados a las fracciones continuadas por dependencia entre ensayos) y pone en duda el caso convergente, asegurando que la probabilidad $a_n < \phi(n)$ se cumpla de alguna n en adelante (significando $a_n < \phi(n), a_{n+1} < \phi(n+1), \ldots$ para alguna n dada) es positivo pero menor que uno. Borel responde en su artículo de 1912 insistiendo que su resultado fue correcto, esto es, A_∞ es cero o uno en el caso que la fracción continuada este de acuerdo con que $\Sigma \frac{1}{\phi(n)}$ sea convergente o divergente. La respuesta dada por Borel en 1912 es la ya expuesta en la seccion 1.3 y 1.4 donde nuevamente situa sus argumentos en la generalidad de su Ley Cero-Uno, es decir, en el espacio de todas las sucesiones numerables de ensayos con dos casos posibles: favorable o desfavorable; mientras que Bernstein se ha concentrado unicamente con las aplicaciones a las fracciones continuadas.

Borel dice que no existe ninguna contradicción entre sus resultados obtenidos en 1909 y los de Bernstein, argumentando la analogía existente entre sus fórmulas (de la (2.9) a la (2.13)) con las obtenidas por Berns_ tein. (de la (2.17) a la (2.21)). La diferencia esencial entre ambos conjuntos de fórmulas es que Bernstein maneja explícitamente la dependencia mientras que Borel no. Es mas, Borel admite que existe dependencia entre ensayos aunque también afirma que no invalida su resultado y que la nueva demostración (sección 1.3) no difiere sensiblemente de la original

La nueva prueba de Borel, ahora si, demuestra los resultados de fracciones continuadas obtenidos en 1909 puesto que explícitamente se está admitiendo la dependencia entre ensayos.

Una observación muy justa que Borel hace en 1912 es que el resultado que Bernstein pone en duda, puede ser deducido de los cálculos mismos de Bernstein; la desigualdad

$$\prod_{\nu=n}^{\infty} \frac{\phi(\nu)-1}{\phi(\nu)+1} < P[a_n < \phi(n), a_{n+1} < \phi(n+1), \dots] < \prod_{\nu=n}^{\infty} \left(1 - \frac{1}{\phi(\nu)}\right)$$

muestra en efecto que si n es tomada bastante grande, los productos infinitos que allí aparecen tienden a uno; que es la misma medida del conjunto de los puntos para los cuales la desigualdad $a_n \geq \phi(n)$ no se verifica para un número finito de valores n, es decir, deja de verificarse a partir de un valor de n suficientemente grande (no fijo de antemano).

CAPITULO 3

Extensión del Lema Borel-Cantelli

El objetivo principal de este capítulo es la presentación de la extensión del Lema de Borel-Cantelli y su uso para obtener los resultados de Borel en lo que respecta a las fracciones continuadas. La exposición que se da de esperanza condicional y martingalas es únicamente la necesaria para llegar a la extensión del Lema de Borel-Cantelli y está tomada de las referencias [3] y [15]. Para una exposición más detallada y completo de estos temas véase las referencias antes citadas.

3.1 Esperanza condicional

Sea $(\Omega, \mathcal{L}, P)$ un espacio de probabilidad, esto es ; Ω un conjunto no vacío, $\mathcal{L}$ un σ-campo de subconjuntos de Ω , P una medida definida en $\mathcal{L}$ que satisface $P(\Omega) = 1$. Sea A un evento fijo en $\mathcal{L}$ tal que $P(A) > 0$. Entonces

$$P\{\cdot \mid A\} = \frac{P(\cdot \cap A)}{(A)}$$

es una medida de probabilidad en $(\Omega, \mathcal{L})$, de modo que la terna $(\Omega, \mathcal{L}, P\{\cdot \mid A\})$ es un espacio de probabilidad.

Sea X una variable aleatoria definida en (Ω, L, P) tal que EX existe. Claramente X es integrable con respecto a $P\{\cdot \mid A\}$ en Ω Definimos,

$$(3.1) \qquad E\{X \mid A\} = \int_{\Omega} X \, dP\{\cdot \mid A\}$$

en en este caso $E\{X \mid A\}$ se conoce como la esperanza condicional de X dado A.

Como

$$P\{\cdot \mid A\} = 0 \qquad \text{es la clase} \qquad \{L \cap A^c : L \in L\}$$

y

$$P\{\cdot \mid A\} = \frac{1}{P(A)} P(\cdot) \qquad \text{es la clase} \qquad \{L \cap A : L \in L\}$$

podemos reescribir (3.1) como

$$(3.2) \quad E\{X \mid A\} = \frac{1}{P(A)} \int_A X \, dP = \frac{1}{P(A)} \quad E(X_{\psi_A})$$

donde X_A es la función indicadora del evento A. En particular para $X = \psi_L$ para $L \in L$, se cumple

$$E\{\psi_L \mid A\} = \frac{P(L \cap A)}{P(A)} = P\{L \mid A\}.$$

Consideremos el σ-campo $M = \{\Omega, A, A^c, \emptyset\}$, donde $\emptyset$ denota el conjunto vacío y A^c el complemento de A, M está generado por $A \in L$. Estableciendo

$$E\{X|M\} = \begin{cases} E\{X \mid A\} & \text{si } w \in A \\ E\{X \mid A^c\} & \text{si } w \in A^c \end{cases}$$

vemos que $E\{X|M\}$ es una variable aleatoria con dos valores, la cual satisface

$$\int_\Omega E\{X|M\}\, dP = P(A)\, E\{X \mid A\} + P(A^c)\, E\{X \mid A^c\}$$

$$= E\{X_{\psi_A}\} + E\{X_{\psi_{A^c}}\}$$

$$= \int_\Omega X\, dP.$$

La función M-medible $E\{X \mid M\}$ definida en Ω se conoce como la esperanza condicional de X, dado el σ-campo M.

Si ahora consideramos el σ-campo $M \subset L$ generado por los conjuntos $A_1, A_2, \ldots, A_n \in L$, para $n \geq 2$, tales que $A_i \cap A_j = \emptyset$ para $i \neq j$, $\bigcup_{i=1}^{n} A_i = \Omega$ y para $P(A_i) > 0$ para $i=1,2,\ldots, n$. Sea X una variable aleatoria tal que EX existe. Entonces definimos

(3.3) $\qquad E\{X|M\} = E\{X \mid A_i\} \quad \text{si } w \in A_i, \quad i=1,2,\ldots, n,$

donde $E\{X|A_i\}$ está definida como en (3.1). La variable aleatoria $E\{X \mid M\}$ es llamada la esperanza de X dado el σ-campo M.

Usando (3.2), obtenemos:

$$P(A_i) \quad E\{X \mid A_i\} = E(X_{\psi_{A_i}}), \qquad i=1,2,\ldots, n$$

de modo que

$$\sum_{i=1}^{n} P(A_i)\, E\{X \mid A_i\} = \sum_{i=1}^{n} E\,(X_{\psi_{A_i}})$$

$$= EX = \int_{\Omega} X\, d\, P.$$

Como $E\{X \mid M\}$ es una variable aleatoria que toma el valor $E\{X \mid A_i\}$ en A_i para $i=1,2,\ldots, n$ y

$$\int_{A_i} E\{X \mid A_i\}\, d\, P = P(A_i)\, E\, \{X \mid A_i\}$$

entonces

$$(3.4) \qquad \int_{\Omega} E\{X \mid M\}\, d\, P = \sum_{i=1}^{n} P(A_i)\, E\,\{X \mid A_i\} = \int_{\Omega} X\, d\, P.$$

Claramente $E\{X|M\}$ es medible con respecto a M . Si X es M medible, entonces $E\{X|M\}= X$ casi dondequiera (c.d.) con respecto a la medida de probabilidad inducida por P en M .

En general, sea $A_1,A_2,\ldots,A_n,\ldots$ una partición de Ω con $P(A_n) > 0$ para $n=1,2,\ldots$ y sea M el σ-campo generado por los conjuntos A_n. Si X es una variable aleatoria tal que EX es finita, entonces

(3.5) $$E\{X \mid \mathcal{M} = \sum_{n=1}^{\infty} E\{X \mid A_n\} \; \psi_{A_n}$$

define la esperanza condicional de X dado $\mathcal{M}$. Asi, $E\{X|\mathcal{M}\}$ es una variable aleatoria discreta que toma los valores $E\{X \mid A_n\}$ para $w \in A_n$ y se sigue que

(3.6) $$\int_{\Omega} E\{X \mid \mathcal{M}\} \, dP = \int_{\Omega} X \, dP.$$

Claramente $E\{X \mid \mathcal{M}\}$ es medible con respecto a $\mathcal{M}$. Si X es $\mathcal{M}$ medible, entonces $E\{X \mid \mathcal{M}\} = X$ casi dondequiera (c.d.) con respecto a la medida de probabilidad inducida por P en $\mathcal{M}$.

En general, sea $A_1, A_2, \ldots A_n, \ldots$ una partición de Ω con $P(A_n) > 0$ para $n=1,2,\ldots$ y sea $\mathcal{M}$ el σ-campo generado por los conjuntos A_n. Si X es una variable aleatoria tal que EX es finita, entonces

(3.5) $$E\{X \mid \mathcal{M}\} = \sum_{n=1}^{\infty} E\{X \mid A_n\} \; \psi_{A_n}$$

define la esperanza condicional de X dado $\mathcal{M}$. Asf, $E\{X|\mathcal{M}\}$ es una variable aleatoria discreta que toma los valores $E\{X \mid A_n\}$ para

$w \in A_n$ y se sigue que

(3.6) $$\int_\Omega E\{X \mid \mathcal{M}\}\, dP = \int_\Omega X\, dP.$$

Claramente $E\{X \mid \mathcal{M}\}$ es medible con respecto a $\mathcal{M}$ y si X es $\mathcal{M}$-medible entonces $E\{X|\mathcal{M}\} = X$ c.d. con respecto a la medida de probabilidad inducida por P en $\mathcal{M}$.

En general, se puede definir la esperanza condicional de una variable aleatoria X dado un σ-campo $\mathcal{D} \subset \mathcal{L}$ como:

Definición 1. Supongamos que X es una variable aleatoria en $(\Omega, \mathcal{L}, P)$ integrable y $\mathcal{D}$ un σ-campo en $\mathcal{L}$. Existe una variable aleatoria $E\{X \mid \mathcal{D}\}$, llamada la esperanza condicional de X dado $\mathcal{D}$, la cual tiene las dos propiedades siguientes:

(i) $E\{X \mid \mathcal{D}\}$ es $\mathcal{D}$-medible e integrable.

(ii) $E\{X \mid \mathcal{D}\}$ satisface la ecuación funcional

(3.7) $$\int_A E\{X \mid \mathcal{D}\}\, dP_{\mathcal{D}} = \int_A X\, dP \qquad A \in \mathcal{D}$$

donde $P_{\mathcal{D}}$ es la restricción de P en $\mathcal{D}$.

Para probar la existencia de tal variable aleatoria, primero consideremos el caso X no negativa. Definimos una medida ν en

$\mathcal{D}$ por

$$\nu(A) = \int_A X \, dP$$

como X es integrable, esta medida es finita y absolutamente continua con respecto a P. Por el teorema de Radón-Nikodym hay una función f, medible en $\mathcal{D}$ tal que

$$\nu(A) = \int_A f \, dP_{\mathcal{D}}$$

Esta f tiene las propiedades (i) y (ii). Si X no necesariamente es positiva, $E(X^+ \mid \mathcal{D}) - E(X^- \mid \mathcal{D})$ tiene las propiedades requeridas, donde

$$X^+(\omega) = \begin{cases} X(\omega) & \text{si } X(\omega) \geq 0 \\ 0 & \text{si } X(\omega) < 0 \end{cases}$$

$$X^-(\omega) = \begin{cases} 0 & \text{si } X(\omega) \geq 0 \\ -X(\omega) & \text{si } X(\omega) < 0 \end{cases}$$

Algunas propiedades elementales de la función $E\{X \mid \mathcal{D}\}$ se enuncian en la siguiente proporción.

<u>Proposición 1</u>. Sea X una variable aleatoria definida en $(\Omega, \mathcal{L}, P)$ tal

que EX existe, y sea $\mathcal{D} \subset \mathcal{L}$ un σ-campo

(i) Si X es $\mathcal{D}$-medible, entonces $E\{X|\mathcal{D}\} = X$ c.d.

(ii) Si $X = \ell$ donde ℓ es una constante, entonces

$$E\{X \mid \mathcal{D}\} = \ell$$

(iii) Si $X \geq 0$ c.d., entonces $E\{X|\mathcal{D}\} \geq 0$ c.d.

(iv) Si Y es otra variable aleatoria en $(\Omega, \mathcal{L}, P)$ tal que EY existe.

Sean $a, b \in \mathbb{R}$. Entonces

$$E\{aX + bY|\mathcal{D}\} = aE\{X|\mathcal{D}\} + bE\{ Y \mid \mathcal{D} \} \quad \text{c.d.}$$

La prueba es inmediata a partir de la definición

<u>Proposición 2</u>. Sea $\mathcal{D}_1$ y $\mathcal{D}_2$ dos sub-σ-campos de $\mathcal{L}$ tales que $\mathcal{D}_1 \subset \mathcal{D}_2$. Sea X una variable aleatoria tal que EX existe. Entonces

$$E\{X \mid \mathcal{D}_1\} = E\{E\{X|\mathcal{D}_1\} \mid \mathcal{D}_2\} = E\{E\{X \mid \mathcal{D}_2\} \mid \mathcal{D}_1\} \quad \text{c.d.}$$

Prueba. Sean $P_{\mathcal{D}_1}$ y $P_{\mathcal{D}_2}$ las restricciones de P en $\mathcal{D}_1$ y $\mathcal{D}_2$ respectivamente. Entonces para $A \in \mathcal{D}$,

$$\begin{aligned}
\int_A E\{E\{X|\mathcal{D}_2\}|\mathcal{D}_1\} d P_{\mathcal{D}_1} &= \int_A E\{X \mid \mathcal{D}_2\} \, dP_{\mathcal{D}_2} \\
&= \int_A X \, dP \\
&= \int_A E\{X \mid \mathcal{D}_1\} \, dP_{\mathcal{D}_1}
\end{aligned}$$

por lo tanto $E\{X \mid \mathcal{D}_1\} = E\{E\{X \mid \mathcal{D}_2\} \mid \mathcal{D}_1\}$ c.d. $(P\mathcal{D}_1)$

Ahora, notemos que $E\{X|\mathcal{D}_1\}$ es $\mathcal{D}_1$-medible, y por consiguiente es $\mathcal{D}_2$-medible; de modo que por la parte (i) de la proposición 1 se sigue que

$$E\{E\{X \mid \mathcal{D}_1\}|\mathcal{D}_2\} = E\{X \mid \mathcal{D}_1\} = E\{E\{X \mid \mathcal{D}_2\}|\mathcal{D}_1\} \quad \text{c.d.} \quad (P_{\mathcal{D}_1}).$$

3.2 Martingalas

Sea (Ω, L, P) un espacio de probabilidad y sea $(T,<)$ un conjunto parcialmente ordenado. Sea $\{\mathcal{D}_t,\ t \in T\}$ una colección de sub-σ-campos de L tales que $\mathcal{D}_s \subset \mathcal{D}_t$ para $s < t$, con $s, t \in T$. Sea $X = \{X_t,\ t \in T\}$ una colección de variables aleatorias definidas en Ω, cada una con esperanza finita.

Definición 2. La clase X se dice que constituye una martingala con respecto a $\{\mathcal{D}_t,\ t \in T\}$ si las siguientes condiciones se cumplen:

(i) Para todo $t \in T$, X_t es $\mathcal{D}_t$-medible

(ii) Para $s, t \in T$, $s < t$, la relación

$$E\{X_t \mid \mathcal{D}_s\} = X_s \quad \text{c.d.} \tag{3.8}$$

se cumple.

La clase X se dice que es una submartingala (supermartingala) si en (3.8) reemplazamos el signo = por el signo $\geq$ ($\leq$).

Claramente, cambiando X_t por $-X_t$ intercambiamos "submartingala" y "supermartingala".

En lo siguiente, sea $T = \{1,2,\ldots\}$ y si $\{X_n\}$ es una martingala con respecto a $\{\mathcal{D}_n\}$, diremos que $\{X_n, \mathcal{D}_n\}$ es una martingala.

<u>Proposición 3</u>. $\{X_n, \mathcal{D}_n\}$ es una martingala (submartingala o supermartingala) sí y sólo sí para toda $n > 1$

(3.9) $$E\{X_n \mid \mathcal{D}_{n-1}\} = (\geq o \leq) \quad X_{n-1} \qquad \text{c.d.}$$

Prueba. Claramente (3.8) $\Longrightarrow$ (3.9). Para $m < n$, puesto que $\mathcal{D}_m \subset \mathcal{D}_{n-1}$

$$\begin{aligned} E\{X_n \mid \mathcal{D}_m\} &= E\{E\{X_n \mid \mathcal{D}_{n-1}\} \mid \mathcal{D}_m\} \\ &= E\{X_{n-1} \mid \mathcal{D}_m\} \\ &\vdots \\ &= E\{X_{m+1} \mid \mathcal{D}_m\} \\ &= X_m \qquad \text{c.d.} \end{aligned}$$

Para el caso de submartingalas o supermartingalas cambiamos el signo = por los signos $\geq$ o $\leq$ respectivamente.

Teorema 1. (Teorema de Convergencia de Submartingalas)

Sea $\{X_n, \mathcal{D}_n\}$ una submartingala. Supóngase que $\limsup_{n \to \infty} E|X_n| < \infty$. Entonces existe una variable aleatoria X la cual es $\sigma(\bigcup_{n=1}^{\infty} \mathcal{D}_n)$-medible tal que $X_n \xrightarrow{c.d.} X$. Además, la desigualdad

$$E|X| \leq \limsup_{n \to \infty} E|X_n| < \infty$$

se cumple.

Aquí $\sigma(\cdot)$ denota el más pequeño σ-campo inducido por $(\cdot)$.

La prueba del teorema anterior está fuera de los objetivos de este trabajo. Para su demostración consúltese las referencias [3] y [15].

3.3 Lema de Borel-Cantelli extendido

Una de las aplicaciones de la teoría de martingalas es la obtención del lema de Borel-Cantelli extendido.

<u>Proposición 4</u>. Sea $\{X_n\}$ una sucesión de variables aleatorias no negativas c.d., uniformemente acotadas definidas en (Ω, L, P), y sea $\{\mathcal{D}_n\}$ una sucesión no decreciente de sub-σ-campos de L tales que X_n es $\mathcal{D}_n$-medible para cada $n=1,2,\ldots$ Entonces las series

$$\sum_{n=1}^{\infty} X_n \quad y \quad \sum_{n=1}^{\infty} E\{X_n \mid \mathcal{D}_{n-1}\}, \qquad \mathcal{D}_0 = \{\emptyset, \Omega\}$$

ambos convergen c.d. o ambos divergen c.d.

Prueba. Sea $Y_k = X_k - E\{X_k \mid \mathcal{D}_{k-1}\}$ para $k \geq 1$, escribamos $S_n = \sum_{k=1}^{n} Y_k$, $n \geq 1$. Entonces $\{S_n, \mathcal{D}_n\}$ es una martingala, ya que

$$\begin{aligned} E\{S_n \mid \mathcal{D}_{n-1}\} &= E\{S_{n-1} + (X_n - E\{X_n \mid \mathcal{D}_{n-1}\}) \mid \mathcal{D}_{n-1}\} \\ &= E\{S_{n-1} \mid \mathcal{D}_{n-1}\} + E\{X_n \mid \mathcal{D}_{n-1}\} - E\{E\{X_n \mid \mathcal{D}_{n-1}\} \mid \mathcal{D}_{n-1}\} \\ &= E\{S_{n-1} \mid \mathcal{D}_{n-1}\} \\ &= S_{n-1} \end{aligned}$$

con $E\{S_n\} = 0$ para toda $n \geq 1$.

Por el teorema de convergencia de submartingala se sigue que $\lim_{n\to\infty} S_n$ existe y es finito c.d., de modo que

$$P\{\lim_{n\to\infty} S_n = \infty\} = P\{\lim_{n\to\infty} S_n = -\infty\} = 0.$$

Si A es un evento en $\mathcal{L}$ y $\mathcal{D}$ es un sub-σ-campo de $\mathcal{L}$, denotamos por

$$P\{A \mid \mathcal{D}\} = E\{\psi_A \mid \mathcal{D}\}$$

Como un corolario de la proposición 4, obtenemos el lema de Borel-Cantelli extenido.

<u>Lema de Borel-Cantelli Extendido</u>. Sea $\{\mathcal{D}_n\}$ una sucesión no decreciente de sub-σ-campos de $\mathcal{L}$, sea $A_n \in \mathcal{D}_n$ para n=1,2,3,... Escribimos $p_1 = P(A_1)$, y para $n \geq 2$

$$p_n = P\{A_n \mid \mathcal{D}_{n-1}\}.$$

Entonces

$$P\{(\limsup_{n\to\infty} A_n) \,\Delta\, (\sum_{n=1}^{\infty} p_n = \infty)\} = 0.$$

En otras palabras,

$$P(\limsup_{n\to\infty} A_n) = 1 \quad \text{si y solo si} \quad P\{\omega : \sum_{n=1}^{\infty} p_n(\omega) = \infty\} = 1,$$

[Aquí $A \Delta B = (A-B) \cup (B-A)$]

Prueba. Tomemos $X_n = \psi_{A_n}$ para $n \geq 1$, entonces $\{X_n\}$ es una sucesión de variables aleatorias no negativas y uniformemente acotadas, donde X_n es $\mathcal{D}_n$-medible para cada n. Se sigue en forma inmediata de la proposición 4 que las series $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} p_n$ ambas convergen c.d. o ambas divergen c.d.

Por otra parte, observemos que

$$\{\omega : \sum_{n=1}^{\infty} X_n(\omega) = \infty\} = \{\omega: \sum_{n=1}^{\infty} \psi_{A_n}(\omega) = \infty\}$$

$$= \limsup_{n\to\infty} A_n$$

Por lo tanto

$$P\{\sum_{n=1}^{\infty} p_n = \infty\} = 1 \quad \text{si y solo si} \quad P(\limsup_n A_n) = 1$$

o equivalentemente

$$P\{(\limsup_{n\to\infty} A_n) \,\Delta\, (\sum_{n=1}^{\infty} p_n = \infty)\} = 0$$

3.4 Aplicación de la extensión del Lema de Borel-Cantelli a sucesiones infinitas de ensayos Bernoulli

En lo que sigue supondremos que Ω es el conjunto de sucesiones de ensayos infinitos de Bernoulli, esto es

$$\Omega = \prod_{i=1}^{\infty} \{0,1\}_i .$$

Sea $\omega \in \Omega$ tal que $\omega = (\xi_1, \xi_2, \ldots)$ y el conjunto $\{i_1, i_2, \ldots, i_n\}$ un subconjunto finito de enteros positivos. Definimos,

$$\pi_{i_1, i_2, \ldots, i_n}(\omega) = (\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_n}) \in \prod_{j=1}^{n} \{0,1\}_j$$

donde $\pi_{i_1, i_2, \ldots, i_n}$ es llamada una proyección de ω sobre $\overset{\infty}{\underset{j=1}{X}} \{0,1\}_j$.

Sean $E_n = \{\omega \in \Omega \mid \pi_n(\omega) = 0\}$ para $n=1,2,\ldots$

una sucesión de conjuntos en Ω y $\mathcal{D}_n$ el σ-campo generado por los conjuntos de la forma

$$A_{\xi_1, \xi_2, \ldots, \xi_n} = \pi^{-1}_{1,2,\ldots,n} (\xi_1, \xi_2, \ldots, \xi_n)$$

Notemos que los 2^n conjuntos $A_{s_1, s_2, \ldots, s_n}$ forman una partición de para toda n.

Si denotamos por L al σ-campo generado por los σ-campos $\mathcal{D}_n$ para $n=1,2,\ldots,$ entonces claramente la sucesión $\{\mathcal{D}_n\}$ es una sucesión no decreciente de sub-σ-campos de L con $E_n \in \mathcal{D}_n$.

Tomemos $$p_2 = P\{E_2 | \mathcal{D}_1\} = E\{\psi_{E_2} \mid \mathcal{D}_1\}$$

donde

$$E\{\psi_{E_1} \mid \mathcal{D}_1\} = \begin{cases} E\{\psi_{E_2} \mid A_0\} & \text{si} \quad \pi_1(\omega) = 0 \\ E\{\psi_{E_2} \mid A_1\} & \text{si} \quad \pi_1(\omega) = 1 \end{cases}$$

con

$$E\{\psi_{E_2} \mid A_0\} = \frac{1}{P(A_0)} \int_{A_0} \psi_{E_2} \, dP = \frac{P(A_0 \quad E_2)}{P(A_0)}$$

$$= P\{E_2 \mid A_0\}$$

de manera similar obtenemos

$$E\{\psi_{E_2} \mid A_1\} = P\{E_2 \mid A_1\}$$

por lo tanto

$$E\{\psi_{E_2} \mid \mathcal{D}_1\} = \begin{cases} P\{E_2 \mid A_0\} & \text{si} \quad \pi_1(\omega) = 0 \\ P\{E_2 \mid A_1\} & \text{si} \quad \pi_1(\omega) = 1 \end{cases}$$

En general

$$p_n = P\{E_n \mid \mathcal{D}_{n-1}\} = E\{\psi_{E_n} \mid \mathcal{D}_{n-1}\}$$

con

$$E\{\psi_{E_n} \mid \mathcal{D}_{n-1}\} = P\{E_n \mid A_{\xi_1, \xi_2, \ldots, \xi_{n-1}}\} \quad \text{si } \pi_{1,\,2,\ldots,\,n}\,(\omega) =$$

$$= (\xi_1, \xi_2, \ldots, \xi_{n-1})$$

Así, las E_n cumplen con todas las hipótesis del lema de Borel-Cantelli y por lo tanto, para estos casos lo podemos aplicar.

En particular, para aplicarlo a las fracciones continuadas, tomemos la doble desigualdad (2.19) y multipliquémosla por ℓ_{n_1}. Nos queda, usando notación moderna

$$\frac{1}{k} < P[a_1 = m_1, \ldots, a_{n-1} = m_{n-1}, a_n \geq 1] < P[a_1 = m_1, \ldots, a_{n-1} = m_{n-1}, a_n \geq k]$$

$$< \frac{2}{k+1} \; P[a_1 = m_1, \ldots, a_{n-1} = m_{n-1}, a_n \geq 1]$$

Ahora, si a esta última doble desigualdad, por aditividad numerable, la sumamos por ejemplo para a_1, desde el valor fijo m_1 hasta infinito, a_2 desde 1 hasta el valor $m_2, \ldots, a_{n-1}$ desde el valor m_{n-1} hasta infinito obtenemos

$$\frac{1}{k} \, P[a_1 \geq m_1, \; a_2 < m_2, \ldots, a_{n-1} \geq m_{n-1}, \; a_n \geq 1] \quad <$$

$$P[a_1 \geq m_1, \; a_2 < m_2, \ldots, a_{n-1} \geq m_{n-1}, \; a_n \geq k] \quad <$$

$$\frac{2}{k+1} \; P[a_1 \geq m_1, \; a_2 < m_2, \ldots, a_{n-1} \geq m_{n-1}, \; a_n \geq 1]$$

dividiendo entre $P[a_1 \geq m_1,\ a_2 < m_2, \ldots,\ a_{n-1} \geq m_{n-1},\ a_n \geq 1]$

obtenemos

$$\frac{1}{k} < P[a_n \geq k \mid a_1 \geq m_1,\ a_2 < m_2, \ldots,\ a_{n-1} \geq m_{n-1}] < \frac{2}{k+1}$$

de manera semejante, para cualesquiera valores $m_1, m_2, \ldots, m_{n-1}$ ya sea que $a_i \geq m_i$ o $a_i < m_i$ para $i=1, \ldots, n-1$ se cumple que

$$\frac{1}{k} < P[a_n \geq k \mid a_1 \overset{<}{\underset{\geq}{\text{ó}}} m_1,\ a_2 \overset{<}{\underset{\geq}{\text{ó}}} m_2, \ldots,\ a_{n-1} \overset{<}{\underset{\geq}{\text{ó}}} m_{n-1}] < \frac{2}{k+1}$$

por lo tanto, si $\phi(n)$ es una función de valores enteros y tomamos como caso favorable

$$a_n(x) \geq \phi(n)$$

donde $x \in [0,1]$, entonces

$$\frac{1}{\phi(n)} < P[a_n(x) \geq \phi(n) \mid a_1(x) \overset{<}{\underset{\geq}{\text{ó}}} \phi(1), \ldots,\ a_{n-1}(x) \overset{<}{\underset{\geq}{\text{ó}}} \phi(n-1)] < \frac{2}{\phi(n)+1}$$

Poniendo esta desigualdad en la notación desarrollada al principio de esta sección, nos queda

$$\frac{1}{\phi(n)} < P\{E_n \mid A_{\xi_1, \xi_2, \ldots, \xi_{n-1}}\} < \frac{2}{\phi(n)+1}$$

y esta doble desigualdad se cumple para cualesquiera de los 2^{n-1} conjuntos $A_{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_{n-1}}$, así que

$$\frac{1}{\phi(n)} < P\{E_n \mid D_{n-1}\} < \frac{2}{\phi(n)+1}$$

Por lo tanto, aplicando el lema de Borel-Cantelli extendido se sigue:

Si $\sum_n \frac{1}{\phi(n)}$ converge, entonces para toda $\omega \in \Omega$, la serie

$$\sum_{n=1}^{\infty} p_n(\omega) \leq \sum_{n=1}^{\infty} \frac{2}{\phi(n)+1} < \infty$$

y por lo tanto se cumple $P\{\limsup_{n\to\infty} E_n\} = 0$

De manera semejante, si $\sum_n \frac{1}{\phi(n)}$ diverge, entonces para toda $\omega \in \Omega$

$$\sum_{n=1}^{\infty} \frac{1}{\phi(n)} < \sum_{n=1}^{\infty} p_n(\omega) = \infty$$

y por lo tanto $P\{\limsup_{n\to\infty} E_n\} = 1.$

Con esto confirmamos el resultado obtenido por Borel de las fracciones continuadas.

APENDICE

Una condición necesaria y suficiente para la convergencia del producto $\prod_{n=1}^{\infty} a_n$, donde $a_n > 0$, es que la serie $\sum_{n=1}^{\infty} \log a_n$ sea convergente. En efecto, como una consecuencia de la continuidad del logaritmo, las sumas parciales $\sum_{k=1}^{n} \log a_k = \log (a_1.a_2 \ldots a_n)$ de esta serie tenderán a un límite definido si y sólo si, los productos parciales $a_1.a_2 \ldots a_n$ tienen un límite positivo.

En base a esta propiedad, podemos establecer los siguientes lemas:

Lema 1 . Sea $a_n = 1 + \alpha_n$ con $a_n > 0$. El producto

$$\prod_{n=1}^{\infty} (1 + \alpha_n)$$

converge a un valor no nulo, si, y sólo si la serie

$$\sum_{n=1}^{\infty} |\alpha_n|$$

converge y ninguno de los factores $(1 + \alpha_n)$ es nulo.

Prueba. Podemos suponer, omitiendo un número finito de factores si es necesario, que $|\alpha_n| < {}^1/_2$ para toda n. Entonces tenemos por el teorema del valor medio

$$\text{(A.1)} \qquad \log(1+h) = \log(1+h) - \log 1 = \frac{h}{1+\theta h} \qquad \text{con } 0<\theta<1$$

Por consiguiente, de (A.1) obtenemos

$$\frac{2}{3}|\alpha_n| \leq \frac{|\alpha_n|}{1+|\alpha_n|} \leq |\log(1+\alpha_n)| = \left|\frac{\alpha_n}{1+\theta\alpha_n}\right| \leq \frac{|\alpha_n|}{1-|\alpha_n|} \leq 2|\alpha_n|$$

y en esta forma la serie $\sum_{n=1}^{\infty} \log(1+\alpha_n)$ converge si, y sólo si $\sum_{n=1}^{\infty} |\alpha_n|$ converge. De donde el resultado se sigue.

<u>Lema 2.</u> Sea $0 \leq \alpha_n < 1$ para $n=1,2,\ldots$ El producto

$$\prod_{n=1}^{\infty} (1-\alpha_n)$$

es nulo, si, y sólo si la serie

$$\sum_{n=1}^{\infty} \alpha_n = +\infty$$

Prueba. De (A.1) obtenemos que

$$\text{(A.2)} \qquad -\frac{\alpha}{1-\alpha_n} \leq \log(1-\alpha_n) \leq -\alpha_n$$

Admitamos primeramente que la sucesión $\{\alpha_n\}$ no tiene como punto límite la unidad, entonces podemos suponer, omitiendo un número finito de términos si fuera necesario, que

$$\alpha_n < 1 - \varepsilon \tag{A.3}$$

para alguna $\varepsilon > 0$ fija, entonces de (A.2) y (A.3) se sigue que

$$-\frac{\alpha_n}{\varepsilon} \leq \log(1-\alpha_n) \leq -\alpha_n$$

De esta última desigualdad se sigue el resultado.

Si la sucesión $\{\alpha_n\}$ tiene como punto límite la unidad, entonces la desigualdad (A.3) se viola para una infinidad de valores de n y por lo tanto la serie $\sum_{n=1}^{\infty} \alpha_n$ diverge si y sólo si $\sum_{n=1}^{\infty} \log(1-\alpha_n)$ diverge. Con lo que se prueba el resultado para todos los casos.

REFERENCIAS

1. Barone, J. y Novikoff, A. "A History of the Axiomatic Formulation of Probability from Boral to Kolmogorov : Part I".

2. Bernstein, F., 1911. "Über eine Anwendung der Mengenlehre auf ein aus der Theorie der säkularen Störungen herrührendes Problem," Mathematishe Annalen, 71, 417 - 439.

3. Billingsley, P., 1979. "Probability and Measure", Wiley, New York.

4. Borel, E., 1903. "Contribution à l'analyse arithmētique du continu," Journal de Mathimatiques Pures et Appliquées, 9, 329 - 375.

5. Borel, E., 1909. "Les Probabilités dénombrables et leurs applications arithmétiques," Rendiconti del Circolo Matematico di Palermo, 27, 247 - 271.

6. Borel, E., 1912. "Sur un probléme de probabilités relatif aux tractions continues," Mathematische Annalen, 72, 578-584.

7. Borel, E., 1926. "Applications a L'Arithmetique et a la Theorie des Fonctions," Traité du Calcul des Probabilités et de ses applications, tome II, fascicule I, Gauthier-Vllars, Paris.

8. Chung, K.L., 1974. "A Course in Probability Theory", Academic Press, New York.

9. Courant, R., John, F., 1974, "Introducción al Cálculo y al Análisis Matemático", Vol. 1, Editorial Limusa, México.

10. Feller, W., 1978, "Introducción a la Teoría de Probabilidades y sus Aplicaciones", Vol. 1 y 2, Editorial Limusa, México.

11. Gnedenko, G., 1969, "The Theory of Probability", Mir Publishers Moscow.

12. Halmos, P.R., 1974, "Measure Theory", Springer-Verlag, New York.

13. Khintchine, A. Ya., "Continued Fractions", P. Noordhoff, Groningen.

14. Kolmogorov, A.N., "Foundations of the Theory of Probabilities", Chelsea, New York.

15. Laha, R.G., Rohatgi, V.K., 1979, "Probability Theory", Wiley, New York.

16. Royden, H.L., 1968, "Real Analysis", Collier Macmillan, New York.

17. Rudin, W., 1976, "Principles of Mathematical Analysis", 3rd. ed., McGraw-Hill, New York.