UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

Una Introducción al Criterio de Kelly Bayesiano

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Actuario

P R E S E N T A:

Emilio Estevez Ibarra

DIRECTOR DE TESIS:
Frank Patrick Murphy Hernández

CIUDAD DE MÉXICO, 2020

Agradecimientos

A mis padres,

Tete, Sofía,

Tíos, primos y abuelita,

Frank,

Y todos mis profesores y compañeros.

Índice general

Índice de figuras

Índice de tablas

Introducción

Un apostador J que conoce la probabilidad de victoria en una serie de apuestas independientes busca elegir una estrategia que maximice su utilidad en función de la proporción de capital que decida apostar. Por ejemplo, se puede apostar siempre la mitad, una décima parte o la proporción que minimiza la varianza de su capital.

El criterio de Kelly consiste en apostar siempre la misma proporción f^* de capital que maximiza el límite del log-rendimiento, es decir, si K_n es el capital de J después de n apuestas y K_0 su capital inicial se busca maximizar $\lim_{N\to\infty} \frac{1}{N} log \frac{K_N}{K_0}$. Como resutado de la ley de los grandes números es equivalente maximizar el límite del log-rendimiento a maximizar $\frac{1}{n}\mathbb{E}[ln\frac{K_n}{K_0}]$ (capítulo dos), así el criterio de Kelly maximiza el valor esperado del log-rendimiento del capital del apostador.

La problemática principal del Criterio de Kelly es conocer la probabilidad de victoria y se vuelve casi imposible llevarlo a cabo en la práctica, por lo cual el objetivo de este trabajo es presentar el criterio de Kelly bayesiano que cumple dos funciones: estimar la probabilidad de victoria y maximizar el valor esperado de logaritmo natural del capital. Es importante destacar que el criterio de Kelly bayesiano satisface estas dos funciones de manera simultánea.

Para resolver lo anterior se supone una Caminata Aleatoria en un Ambiente Aleatorio (**RWIRE** por sus siglas en inglés), que es una dupla formada por una caminata aleatoria $\{S_n\}$ y sus probabilidades de transición $\{\Delta_n\}_{n=0}^{\infty}$ que es una sucesión de variables aleatorias independientes idénticamente distribuídas tales que $0 < \Delta_n < 1$ para toda n. Para trabajar con una RWIRE se toma una realización $\overline{\alpha} = \{\alpha_n\}_{n=0}^{\infty}$ de $\{\Delta_n\}_{n=0}^{\infty}$ y se procede con los supuestos de una caminata aleatoria con probabilidades de transición $\{\alpha_n\}_{n=0}^{\infty}$. A dicha realización se le conocerá coloquialmente como «fijar» un ambiente y se referirá a $\overline{\alpha}$ como el ambiente «fijo» de la RWIRE.

Para el caso en que el ambiente «fijo» es constante, es decir $\Delta_n = \alpha$ para cada n, se obtiene que la probabilidad de transición de la caminata aleatoria está dada por $\mathbb{P}[S_{n+1} = s_n + 1 | S_n = s_n] = \mathbb{E}[\alpha | S_n = s_n]$ (capítulo tres), que se puede calcular suponiendo una función de densidad *a priori* h para α.

El criterio de Kelly bayesiano consiste en calcular las proporciones de capital f_j que se tienen que apostar para maximizar $\mathbb{E}[ln(K_n)]$ suponiendo un ambiente «fijo» constante y una función de densidad *a priori* para α.

En este trabajo se introducirán los conceptos necesarios para desarrollar el criterio de Kelly bayesiano y se aplicará en el caso particular en que la distribución *a priori* de α es beta.

En el capítulo uno se describirán los antecedentes del criterio de Kelly no bayesiano: la teoría de la información desarrollada por Claude Shannon, el artículo donde John Kelly describe por primera vez el criterio que lleva su nombre y la incursión de Edward Thorp en las apuestas deportivas, el BlackJack y la bolsa de valores.

En el capítulo dos se desarrollará de manera detallada el criterio de Kelly no bayesiano cuando se tiene un canal binario de comunicación y se compararán la tasa de crecimiento exponencial con la tasa de transmisión de comunicación definida por Shannon. Además se generalizará el criterio de Kelly no bayesiano cuando se tiene un canal n-ario de comunicación.

En el capítulo tres se definen los conceptos de RWIRE y variables aleatorias intercambiables y se enunciará el teorema de De Finetti, seguido del desarrollo del criterio de Kelly bayesiano cuando se fija un ambiente constante. El caso en que la distribución *a priori* del ambiente fijo es beta se trabajará con detalle.

En el capítulo cuatro se realizará una simulación Monte Carlo sobre el MAPE (*Mean Absolute Percentage Error*) entre los logaritmos naturales del capital de J utilizando el criterio de Kely no bayesiano y el criterio de Kelly bayesiano.

Finalmente, en el capítulo cinco, se tomarán los resultados de temporadas regulares de equipos de la *NFL*, *NBA* y *MLB* y con base en ellos se simulará el capital de J utilizando el criterio de Kelly bayesiano suponiendo una distribución a *a priori* beta.

Capítulo 1

Antecedentes

Imagine un juego en el cual ganará, o perderá, tanto dinero como decida apostar al resultado de lanzar una moneda. Además, si se supone que la moneda no es justa y conoce la probabilidad, $p > 0.5$, de que el volado favorezca una de las caras de la moneda, es natural que un jugador J que decide participar en este juego se pregunte por la estrategia que resulte en las mejores ganancias.

Una estrategia es apostar la totalidad del capital a la cara que favorece la moneda y así obtener las mayores ganancias lo más rápido posible. Si K_0 y K_i denotan el capital inicial y el capital después de la i-ésima apuesta de J, respectivamente, después de un juego victorioso el capital se duplicaría, una segunda victoria significaría cuadruplicar el capital inicial y de manera general, después de una racha de N juegos victorioso el capital sería $K_N = 2^N K_0$. Sin embargo, en cada juego se apuesta todo el capital y la probabilidad de perderlo después de una racha ganadora de N juegos es $q = 1 - p^N$ y mientras el número de apuestas crece esta probabilidad tiende a uno y la ruina es segura.

Después de analizar el caso anterior uno pensaría en que ser conservador. No apostar o apostar una pequeña fracción del capital parece ser la respuesta. Sin embargo esta estrategia resulta en ganancias nulas o que crecen muy lentamente y podrían no ser competentes con las expectativas de J.

Así, J puede tomar tantas estrategias como se le ocurran, puede apostar en cada juego la mitad, una décima parte o una fracción que minimice la varianza de su capital. Una de las tantas estrategias consiste en maximizar el valor esperado de la utilidad del capital de J, en particular cuando se maximiza el valor esperado del log-rendimiento del capital se obtiene la estrategia conocida como el criterio de Kelly.

El criterio de Kelly se formuló con base en ideas de la teoría de la información creada por Claude Eldwood Shannon (1916 - 2001), matemático, ingeniero eléctrico, criptógrafo estadounidense y compañero de trabajo de John Kelly. Shannon es conocido como el «padre de la teoría de la información» por su artículo «*A mathematical theory of communication*» [**S**].

Shannon consideró un sistema de comunicación [**S**, pp. 2] y planteó la problemática de transmitir de manera exacta, o muy aproximada, el mensaje original hacia su destino suponiendo la existencia de ruido en el canal de transmisión (distorsión del mensaje). Si el receptor conoce las probabilidades de transmisión $p_1, \ldots, p_n$ de n mensajes posibles y recibió el mensaje i-ésimo podría medir lo incierto que es dicho mensaje. La métrica que definió Shannon para medir la incertidumbre es la entropía, H [**S**, pp. 11], de las probabilidades $p_1, \ldots, p_n$, término que, por su definición, se asemeja al utilizado en estadística mecánica. Por ejemplo, H es la H en el famoso teorema H de Boltzman.

Shannon había creado una nueva teoría y demostrado lo necesario para impulsar décadas de desarrollo en el campo de las comunicaciones, muchos de los avances tecnológicos de la segunda mitad del siglo pasado se deben a Shannon. *A mathematical theory of communication* fue un trabajo brillante y sin precedentes.

La teoría de la información funcionaba a la perfección en el campo de las comunicaciones, Shannon había desarrollado tanta teoría que muchos empezaban a buscar aplicaciones en otros campos. Un doctor en física lo haría.

John Larry Kelly Jr. (1923-1965) fue un físico y computológo estadounidense que trabajó para la compañía de investigación *Bell Laboratories*. En la década de los 50's *Bell laboratories* era considerado como uno de los centros de investigación más prestigiosos de Estados Unidos, principalmente por el hecho de que Shannon trabajaba ahí. Se decía que *Bell Labs* era como una universidad en la que los investigadores no daban clase y tenían siempre dinero suficiente para realizar experimentos. En el ranking de genios de Bell Laboratories, Shannon se encontraba en la primera posición y el consenso general ubicaba a Kelly en segundo lugar [**P**, pp. 63]. Shannon y Kelly no trabajaban en el mismo departamento y no se conocían en persona. Sin embargo Kelly era consciente del trabajo de Shannon y había desarrollado un peculiar trabajo que reinterpretaba por completo la teoría de la comunicación. Kelly buscó y se reunió con Shannon para exponerle su artículo de apuestas «*Information Theory* *and Gambling*», Shannon instó a Kelly a publicarlo [**P**, pp. 67].

Kelly define en su trabajo la tasa exponencial de crecimiento del capital $G = \lim_{N\to\infty} \frac{1}{N} log \frac{K_N}{K_0}$ donde *log* es el logaritmo en base dos, su objetivo fue encontrar la proporción de capital que maximizara G.

Volviendo al juego de la moneda no justa y suponiendo que J decide apostar a la cara que favorece la moneda un porcentaje $0 < l < 1$ de su capital, después de N juegos su capital será $K_N = K_0(1 + l)^{W(N)}(1 - l)^{L(N)}$ donde $W(N)$ y $L(N)$ denotan el número de victorias y derrotas después de N apuestas, respectivamente. Se puede demostrar que para este ejemplo que $G = plog(1 + l) + (1 - p)log(1 - l)$. Lo siguiente que plantea Kelly es maximizar G respecto a l. El resultado que obtiene es inusual, la tasa máxima de crecimiento exponencial del capital es igual a la tasa de transmisión de información que define Shannon, $G_{max} = R$ [**P**, pp. 73].

Para el juego de la moneda la proporción que maximia G resulta ser $l = 2p - 1$. El criterio de Kelly indica que J debe apostar cuando $l > 0$ y la proporción del capital a apostar es l. Entonces J debe apostar una proporción equivalente a $2p - 1$ de su capital en el juego de la moneda a la cara favorecida por la moneda.

El artículo termino llamándose «*A new interpretation of information rate*» a petición de *AT&T* que temía que el título original levantara sospechas de que *Bell Laboratories* hiciera trabajos para los apostadores ilegales [**P**, pp. 76].

Kelly moriría 9 años después de publicar «*A new interpretation of information rate*», sin haber llevado a la práctica su propio criterio, sin embargo, su trabajo no pasaría despercibido, el matemático conocido como el padre de la teoría del conteo de cartas se encargaría de probarlo y popularizarlo.

Edward Oakley Thorp (1932-) había empezado a trabajar como profesor de matemáticas en 1959 a la par de estar decidido de vencer a la casa en el juego de cartas *BlackJack* (también conocido como 21). Tras una desastrosa visita a los casinos de Las Vegas para jugar *BlackJack*, basándose en un sistema deficiente, se convenció de encontrar una estrategia óptima para jugar *BlackJack*. Por las grandes posibilidades de manos que hay en un juego de *BlackJack* utilizó la computadora pricipal del MIT para encontrar un método significativamente mejor al que había usado antes.

Hecho ésto, Thorp buscó publicar su método en la revista científica más prestigiosa a su parecer: *The Proceedings of the National Academy of Sciences*. Para publicar en dicha revista científica se requería de un miembro de la academia nacional. El único miembro, matemático, trabajando en el MIT era Shannon. Thorp se reunió con Shannon y le expuso su método. «*A favorable strategy for twenty one*» se publicó en enero de 1961 en *The Proceedings of the National Academy of Sciences*.[**P**, pp. 81].

Thorp y Shannon se embarcarían a los casinos de Las Vegas a poner en práctica la estrategia descrita por Thorp y una pequeña computadora inventada por ambos la cual predecía con cierto margen los resultados de la ruleta [**P**, pp. 104]. La única preocupación de Thorp era la catidad de dinero a apostar en cada apuesta; cualquier jugador, bueno o malo, puede irse a la ruina si no administra bien su capital. Shannon no compartía la misma preocupación dado su conocimiento del criterio de Kelly, el cual propuso como estrategia. En compañía de sus esposas, Vivian Thorp y Betty Shannon, los viajes a Las Vegas fueron fructíferos demostrando la eficiencia del método de Thorp y más importante, el criterio de Kelly.

Thorp se convertiría en una celebridad en 1961 después de publicar *Beat the Dealer*, el best-seller, que explicaba el método de conteo de cartas. El *BlackJack* fue el inicio para Thorp; utilizó el criterio de Kelly en toda clase de apuestas y juegos, desde apuestas deportivas hasta el mercado de valores.

Las apuestas deportivas fueron presa fácil del criterio de Kelly cuando Thorp se alió a un joven y talentoso doctor en ciencias de la computación que había desarrollado un eficiente sistema de apuestas. Una inversión inicial de 50,000 dólares terminó en 123,000 dólares después de apostar 101 días [**T**, pp 18]. Thorp teminaría abandonando las apuestas deportivas a pesar de lo positivos que parecían ser los resultados, pues no cumplían con sus expectativas económicas. Iniciría entonces su incursión en el mercado de valores de *Wall Street*.

El mercado de valores dejó las más grandes ganancias tanto para Thorp como para los inversionistas que ponían en sus manos la administración de su capital. En el mercado de valores se puede diversificar la inversión con la intención de reducir el riesgo y buscar mejores oportunidades de crecimiento de capital. Un inversionista puede invertir su capital en una cantidad n de instrumentos financieros colocando en cada uno de ellos una proporción k_i del total de su capital. Se puede calcular la media y desviación estándar del portafolio en función de la proporción invertida en cada instrumento así como las k_i para minimizar varianzas, método que a Markowitz le valió el premio Nobel. Una manera interesante de invertir es buscando aplicar el criterio de Kelly a la colección de instrumentos, conocida como portafolio, en los que se coloca el capital. De esta forma Thorp obtuvo las mejores ganancias imaginables. Aplicando el criterio de Kelly a un portafolio de activos durante 20 años se obtuvo una tasa de rendimiento de 20 % y una desviación estándar del 6 % contra un promedio de +4.2 % del *Dow Jones* [**T**, pp 35]. El mismo Thorp estima que el valor de las ganancias que hizo para sus inversores es de 80 mil millones de dólares, un promedio de un millón 250 mil apuestas con un valor de 65,000 dólares cada una [**T**, pp 36].

El criterio de Kelly en su planteamiento busca maximizar el límite del log-rendimiento cuando el número de apuestas tiende a infinito. Un millón 250 mil apuestas en 20 años, para efectos prácticos, es lo suficientemente grande para demostrar la eficiencia del criterio y resulta ser una atractiva estrategia para inversores que piensan en apuestas a largo plazo.

Actualmente el criterio de Kelly sigue siendo investigado, se plantean situaciones más generales respecto a la distribución de las variables aleatorias que modelan las apuestas, sean variables conocidas o no, y la amplia gama de herramientas estadísticas, en conjunto con un creciente poder computacional, busca escenarios donde se pueda implementar el criterio de Kelly.

Capítulo 2

Criterio de Kelly no bayesiano

1. Canal binario de comunicación

Primero se considera un canal binario de comunicación, es decir, un canal donde solo se pueden transmitir dos caracteres, por ejemplo, el resultado de la final del mundial de futbol, el *super bowl* o la final en *Wimbledon*. El carácter se transmite de manera correcta con probabilidad p y de manera errónea con propabilidad $q = 1 - p$. Cuando $p = 1$ o $p = 0$ se dice que el canal binario de comunicación no tiene ruido, en caso contrario se dice que el canal binario de comunicación tiene ruido.

Cuando se tiene un canal binario de comunicación sin ruido un apostador, J, que puede conseguir apuestas regularmente debería apostar la totalidad de su capital en cada apuesta, pues ganará con certeza. Sean K_0 su capital inicial y K_N su capital después de N apuestas. En cada apuesta duplicará su capital y al término de N juegos $K_N = 2^N K_0$.

Kelly define la tasa de crecimiento exponencial del apostador J [**K**, pp 919], G, como:

$$G = \lim_{N\to\infty} \frac{1}{N} log \left(\frac{K_N}{K_0} \right)$$

donde $log(\cdot)$ es la función logaritmo en base dos.

Para el caso de una canal binario de comunicación sin ruido se tiene que

$$\begin{aligned} G &= \lim_{N\to\infty} \frac{1}{N} log\left(\frac{2^N K_0}{K_0}\right) \\ &= \lim_{N\to\infty} \frac{1}{N} log\left(2^N\right) \\ &= \lim_{N\to\infty} log\, 2 \\ &=1. \end{aligned}$$

La tasa de crecimiento exponencial del apostador para el caso sin ruido es uno.

Si el canal binario de comunicación fuera ruidoso, J no debería seguir la misma estrategia y apostar la totalidad de su capital en cada apuesta. En este caso la probilidad de ganar N juegos consecutivos es p^N y, en consecuencia, la propabilidad de perder después de una racha ganadora de N juegos es $1 - p^N$ que tiende a uno cuando N tiende a infinito y un jugador que apueste la totalidad de su capital en cada juego lo perdería con probabilidad uno. Se supone entonces que J decide apostar una fracción de su capital l, $0 < l < 1$, en cada apuesta. Así el capital K_{i+1} de J después de la $i+1$-ésima apuesta es $K_i(1+l)$ con probabilidad p si J ganó el i-ésimo juego o $K_i(1-l)$ con probabilidad q si J perdió el i-ésimo juego. De esta manera se tiene que

$$K_N = K_0(1+l)^{W(N)}(1-l)^{L(N)}$$

donde $W(N)$ y $L(N)$ denotan el número de victorias y derrotas de J después de N apuestas, respectivamente. Se observa lo siguiente

- $N = W(N) + L(N)$.
- Sean $\{X_n\}_{n\geq 1}$ variables aleatorias independientes e identicamente distribuidas $Blli(p)$ donde X_i es la variable aleatoria de victoria de la i-ésima apuesta. Note que $W(N) = \sum_{i=1}^{N} X_i$, es decir $W(N) \sim Bin(N,p)$.
- De manera análoga, si $\{Y_n\}_{n\geq 1}$ son variables aleatorias independientes e identicamente distribuidas $Blli(q)$ donde Y_i es la variable aleatoria de derrota de la i-ésima apuesta entonces $L(N) = \sum_{i=1}^{N} Y_i$, es decir $L(N) \sim Bin(N,q)$.

La tasa de crecimiento exponencial para la estrategia l, denotada por $G(l)$, es

$$\begin{aligned} G(l) &= \lim_{N\to\infty} \frac{1}{N} log\left(\frac{K_0(1+l)^{W(N)}(1-l)^{L(N)}}{K_0}\right) \\ &= \lim_{N\to\infty} \frac{1}{N}\left(W(N)log(1+l) + L(N)log(1-l)\right) \\ &= \lim_{N\to\infty} \frac{W(N)}{N} log(1+l) + \lim_{N\to\infty} \frac{L(N)}{N} log(1-l) \end{aligned}$$

después, por la *ley fuerte de los grandes números* (ver apéndice [Ley fuerte de los grandes números])

$$G(l) = p\, log(1+l) + q\, log(1-l)$$

casi seguramente.

PROPOSICIÓN 1. $G(l) = \frac{1}{n}\mathbb{E}\left[ln\left(\frac{K_n}{K_0}\right)\right]$

Dem.

$$\begin{aligned} \mathbb{E}\left[ln\left(\frac{K_n}{K_0}\right)\right] &= \mathbb{E}[ln\left((1+l)^{W(n)}(1-l)^{L(n)}\right)] \\ &= \mathbb{E}[ln\left((1+l)^{W(n)}\right)] + \mathbb{E}[ln\left((1-l)^{L(n)}\right)] \\ &= \mathbb{E}[W(n)ln(1+l)] + \mathbb{E}[L(n)ln(1-l)] \\ &= ln(1+l)\mathbb{E}[W(n)] + ln(1-l)\mathbb{E}[L(n)] \\ &= n\, p\, ln(1+l) + n\, q\, ln(1-l) \end{aligned}$$

□

Por la proposición anterior es equivalente maximizar $G(l)$ respecto a l a $\mathbb{E}[ln\frac{K_n}{K_0}]$ respecto a l, para esto, se hace uso de la siguiente proposición.

PROPOSICIÓN 2. *Sean $n \in \mathbb{N}$, $a = (a_1, \ldots, a_n) \in \mathbb{R}^n_{>0}$, $X \in \mathbb{R}$ y $Z\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por $Z(x_1, \ldots, x_n) = \sum_{i=1}^n a_i log(x_i)$, restringida a $\sum_{i=1}^n x_i = X$. Entonces el valor máximo de Z es alcanzado en x^*, cuya i-ésima entrada está dada por*

$$x_i^* = \frac{X}{A}a_i \quad con\, 1 \le i \le n \quad donde\, A = \sum_{i=1}^n a_i$$

*Una demostración de la proposición puede verse en [**C**].* □

Así las cosas, el punto x^* que maximiza G es

$$x^* = (2p, 2q)$$

Ya que $A = p + q = 1$ y $X = 1 + l + (1 - l) = 2$. Después

$$x_1^* = 1 + l = 2p,$$

$$x_2^* = 1 - l = 2q,$$

entonces

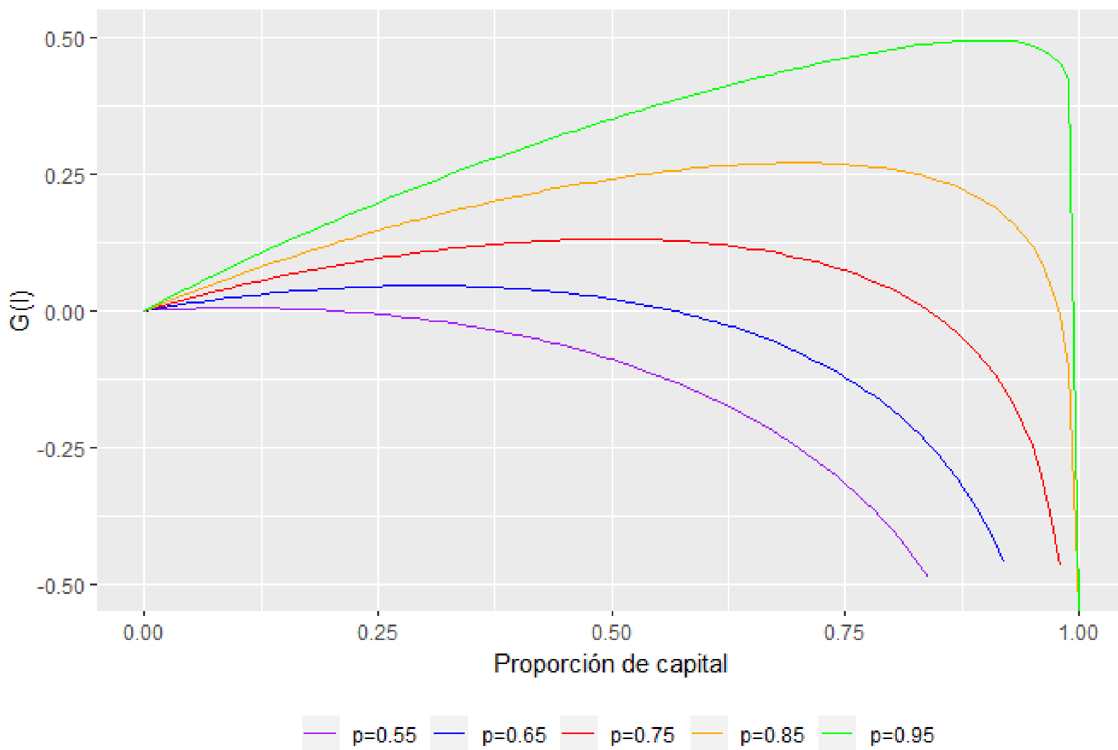
$$l^* = 2p - 1 = p - q$$

finalmente, evaluando l^* en $G(l)$

$$\begin{aligned} G(l^*) =& p\,log(1 + (p - q)) + q\,log(1 - (p - q)) \\ =& p\,log((1 - q) + p) + q\,log((1 - p) + q) \\ =& p\,log(2p) + q\,log(2q) \\ =& p(log(2) + log(p)) + q(log(2) + log(q)) \\ =& p + p\,log(p) + q + q\,log(q) \\ =& 1 + p\,log(p) + q\,log(q). \end{aligned}$$

Considere la figura uno donde se grafica $G(l)$ contra la proporción de capital que J apuesta.

FIGURA 1. Tasa de crecimiento exponencial de J con distintas probabilidades de victoria



Se observa lo siguiente: (i) $G(l)$ alcanza su máximo en $l^* = 2p - 1$ para cada p; (ii) l^* como función de p es creciente, así la utilidad es mayor cuando la probabilidad de victoria es mayor; (iii) $\lim_{l\to 0} G(l) = 0$ para cada p, es decir, cuando J apuesta una pequeña fracción de su capital en todas las apuestas su utilidad es nula y su capital inicial permanecerá casi constante mientras apueste, y (iv) $\lim_{l\to 1^-} G(l) = -\infty$ para cada p, es decir, cuando J apuesta una gran fracción de su capital en cada apuesta su utilidad decrece y perderá casi la totalidad de su capital mientras apueste.

Se puede interpretar a G como una «tasa». Considere i una tasa de interés capitalizable cada apuesta, K_0 el capital inicial a invertir y K_N el capital después de N apuestas. Se tiene lo siguiente

$$K_N = K_0(1+i)^N \quad \Rightarrow \quad \frac{K_N}{K_0} = (1+i)^N$$

sustituyendo $(1+i)^N$ por $\frac{K_N}{K_0}$ en la definición de G se tiene

$$\begin{aligned}
G &= \lim_{N\to\infty} \frac{1}{N} log(1+i)^N \\
&= \lim_{N\to\infty} \left(log(1+i)^N\right)^{\frac{1}{N}} \\
&= \lim_{N\to\infty} log(1+i) \\
&= log(1+i).
\end{aligned}$$

Como la función log es en base 2

$$2^G = 2^{log(1+i)} = 1+i \quad \Rightarrow \quad 2^G - 1 = i$$

Para el caso de un canal binario de comunicación sin ruido se obtuvo $G=1$, entonces $i = 2^1 - 1 = 1$, Esto quiere decir que el capital de J crece de manera equivalente a una inversión con tasa de interés compuesto del 100 % capitalizable cada apuesta.

Para el caso de una canal binario de comunicación con ruido

$$\begin{aligned}
i =& 2^G - 1 \\
=& 2^{1+p\,log(p)+q\,log(q)} - 1 \\
=& 2\left(2^{p\,log(p)}\right)\left(2^{q\,log(q)}\right) - 1 \\
=& 2\left(2^{log(p^p)}\right)\left(2^{log(q^q)}\right) - 1 \\
=& 2p^p q^q - 1.
\end{aligned}$$

Es decir, el capital de J crece de manera equivalente a una inversión con tasa de interés compuesto de $2p^pq^q - 1$ capitalizable cada apuesta. Así, es equivalente para J realizar apuestas de proporción l^* a invertir con una tasa de interés compuesto $i = 2p^pq^q - 1$ capitalizable cada apuesta.

Ahora se compararán los resultados a los que Shannon y Kelly llegaron en sus artículos.

Sean X la variable aleatoria que denota el carácter transmitido en el canal binario de comunicación e Y la variable aleatoria que denota el carácter recibido en el canal binario de comunicación. La entropía de la variable aleatoria X definida por Shannon

[**S**, pp. 11] está dada por

$$H(X) = -\sum_{i=0}^{1} p_i \, log(p_i),$$

donde p_i denota la probabilidad de transmisión del i-ésimo carácter.

Se supone que $X \sim Blli(\frac{1}{2})$, de esta manera se tiene

$$\begin{aligned} H(X) &= -\sum_{i=0}^{1} p_i \, log(p_i) \\ &= -\left(\frac{1}{2} log\left(\frac{1}{2}\right) + \frac{1}{2} log\left(\frac{1}{2}\right)\right) \\ &= -\left(\frac{1}{2}(-1) + \frac{1}{2}(-1)\right) \\ &=1. \end{aligned}$$

Es decir, la entropía de X es uno cuando los caracteres se transmiten con igual probabilidad.

Por otro lado, se supone que $\mathbb{P}(Y = 0 \mid X = 0) = \mathbb{P}(Y = 1 \mid X = 1) = p$ y $\mathbb{P}(Y = 0 \mid X = 1) = \mathbb{P}(Y = 1 \mid X = 0) = q$, es decir, la probabilidad de que que el carácter se haya transmitido de manera correcta es p y de que se haya transmitido de manera incorrecta es q. Después, por el *teorema de la probabilidad total* [**R1**, pp. 77] se tiene lo siguiente

$$\begin{aligned} \mathbb{P}(Y = 0) =& \mathbb{P}(Y = 0 \mid X = 0)\mathbb{P}(X = 0) + \mathbb{P}(Y = 0 \mid X = 1)\mathbb{P}(X = 1) \\ =& \frac{1}{2}p + \frac{1}{2}q = \frac{1}{2}. \end{aligned}$$

De lo que se sigue que $\mathbb{P}(Y = 1) = \frac{1}{2}$. Entonces $Y \sim Blli(\frac{1}{2})$.

La entropía condicional de X dado que se conoce Y se define como sigue [**S**, pp. 12]

$$H_Y(X) = -\sum_{i=0}^{1}\sum_{j=0}^{1} p(i,j) log\left(p_j(i)\right),$$

donde $p(i,j) = \mathbb{P}(X = i, Y = j)$ y $p_j(i) = \mathbb{P}(X = i \mid Y = j)$.

Se procede a calcular $\mathbb{P}(X = i, Y = j)$ y $\mathbb{P}(X = i \mid Y = j)$ para $i, j \in \{0, 1\}$.

Por la definición de probabilidad condicional se tiene que $\mathbb{P}(X = i, Y = j) = \mathbb{P}(Y = j \mid X = i)\mathbb{P}(X = i)$ y con los supuestos anteriores se tiene que $\mathbb{P}(X = 0, Y = 0) = \frac{1}{2}p, \mathbb{P}(X = 0, Y = 1) = \frac{1}{2}q, \mathbb{P}(X = 1, Y = 0) = \frac{1}{2}q, \mathbb{P}(X = 1, Y = 1) = \frac{1}{2}p$.

Después, por el *teorema de Bayes* [**R1**, pp. 84]

$$\mathbb{P}(X = i \mid Y = j) = \frac{\mathbb{P}(Y = j \mid X = i)\mathbb{P}(i = 0)}{\mathbb{P}(Y = j)}.$$

Entonces $\mathbb{P}(X = 0 \mid Y = 0) = p, \mathbb{P}(X = 0 \mid Y = 1) = q, \mathbb{P}(X = 1 \mid Y = 0) = q$ y $\mathbb{P}(X = 1 \mid Y = 1) = p$, y se tiene que

$$\begin{aligned}
H_Y(X) &= -\sum_{i=0}^{1}\sum_{j=0}^{1} p(i,j)log\left(p_j(i)\right) \\
&= -\left(p(0,0)log\left(p_0(0)\right) + p(0,1)log\left(p_1(0)\right) + p(1,0)log\left(p_0(1)\right) + p(1,1)log\left(p_1(1)\right)\right) \\
&= -\left(\left(\frac{p}{2}\right)log(p) + \left(\frac{q}{2}\right)log(q) + \left(\frac{q}{2}\right)log(q) + \left(\frac{p}{2}\right)log(p)\right) \\
&= -\left(p\,log(p) + q\,log(q)\right).
\end{aligned}$$

Finalmente, Shannon define la tasa de transmisión de la información [**S**, pp. 20] como sigue

$$R = H(X) - H_Y(X),$$

y se comprueba que la tasa de crecimiento exponencial definida por Kelly es igual a la tasa de transmisión de información definida por Shannon para el caso de un canal biario de comunicación

$$\begin{aligned}
R =&H(X) - H_Y(X) \\
=&1 - (-p\,log(p) - q\,log(q)) \\
=&1 + p\,log(p) + q\,log(q).
\end{aligned}$$

Para concluir la sección se analiza la estrategia de J de acuerdo al criterio de Kelly.

La fracción de capital que maximiza G es $2p - 1$. Si $p \geq 1/2$ entonces $0 \leq l \leq 1$ y esa es la proporción de capital que J debe apostar.

Por otro lado, si $p \leq 1/2$ entonces $-1 \leq l \leq 0$ y se tiene que apostar una proporción negativa del capital, esto es apostar la fracción de capital $-l$ en contra del resultado cuya probabilidad es p. De no estar permitido apostar en contra el criterio de Kelly dicta que la proporción de capital a apostar es $f^* = max\{l, 0\}$.

2. Canal n-ario

Se considera ahora el caso en que el canal de comunicación puede recibir n caracteres distintos, no necesariamente con igual probabilidad, que representan el resultado de eventos aleatorios. Se supone además que los caracteres son transmitidos de manera independiente. En lo que sigue se utilizará la notación de Kelly [**K**, pp 920]

$p(s)$**:** la probabilidad de que el s-ésimo carácter se haya transmitido.

$p(r \mid s)$**:** la probabilidad condicional de recibir el r-ésimo carácter dado que se transmitió el s-ésimo carácter.

$p(s, r)$**:** la probabilidad conjunta de transmitir el s-ésimo carácter y recibir el r-ésimo.

$q(r)$**:** la probabilidad de recibir el r-ésimo carácter.

$q(s \mid r)$**:** la probabilidad condicional de transmitir el s-ésimo carácter dado que se recibió el r-ésimo carácter.

α_s**:** la proporción pagada en la ocurrencia del s-ésimo carácter transmitido, es decir, α_s es el número de unidades monetarios que se pagan por una apuesta de una unidad monetaria (incluída la unidad apostada).

$a(s \mid r)$**:** la fracción de capital que el apostador decide apostar en la ocurrencia del s-ésimo carácter dado que se recibió el r-ésimo carácter.

Sean un apostador, J, y un canal de comunicación n-ario que puede transmitir de entre un total de n caracteres de un conjunto C de caracteres. J puede realizar apuestas a cualquiera de los n resultados de manera simultánea, a pesar de que solo uno será el ganador. Primero se analizará el caso en el que las proporciones de pago son justas, es decir, las proporciones de pago y las probabilidades de ocurrencia mantienen una relación inversamente proporcional

$$\alpha_s = \frac{1}{p(s)}, \quad \forall s \in C.$$

Sin pérdida de generalidad se supone que J apostará la totalidad de su capital independientemente del carácter que se reciba, es decir

$$\sum_{s \in C} a(s \mid r) = 1, \quad \forall r \in C.$$

Como se suponen proporciones de pago justas, el apostador puede, de manera efectiva, conservar parte de su capital realizando apuestas que se cancelen mutuamente.

Sean K_0 el capital inicial del apostador, K_N el capital del apostador después de N apuestas y $W_{sr}(N)$ el número de veces en las que el carácter transmitido fue s y el carácter recibido fue r, con $s, r \in C$. Se nota que si K_i es el capital de J después de la i-ésima apuesta y suponiendo que s es el carácter transmitido y h es el carácter recibido, con $s, h \in C$ fijos, entonces

$$K_{i+1} = a(s \mid h)\alpha_s K_i$$

dado que J puede apostar simultáneamente en varios resultados, el capital después de la i-ésima apuesta también se puede escribir como

$$K_{i+1} = \prod_{x,r \in C} [a(x \mid r)\alpha_x]^{W_{xr}(i+1)} K_i$$

de está manera

$$K_N = \prod_{x,r \in C} [a(x \mid r)\alpha_x]^{W_{xr}(N)} K_0.$$

Se observa que, para cada $x, r \in C$, la sucesión de variables aleatorias $\{X_n^{xr}\}_{n \in \mathbb{N}}$ es tal que, para cada $n \in \mathbb{N}$, X_n^{xr} se distribuye $Blli(p(x, r))$. Como las variables son independientes e idénticamente distribuidas se tiene que

$$W_{xr}(N) = \sum_{i=1}^{N} X_n^{xr}$$

entonces, $W_{xr}(N) \sim Bin(N, p(x,r))$.

Se procede a calcular la tasa G

$$
\begin{aligned}
G &= \lim_{N\to\infty} \frac{1}{N} log\left(\frac{K_N}{K_0}\right) \\
&= \lim_{N\to\infty} \frac{1}{N} log\left(\frac{\prod_{x,r\in C}[a(x\mid r)\alpha_x]^{W_{xr}(N)}V_0}{V_0}\right) \\
&= \lim_{N\to\infty} \frac{1}{N} log\left(\prod_{x,r\in C}[a(x\mid r)\alpha_x]^{W_{xr}(N)}\right) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{x,r\in C} log\left([a(x\mid r)\alpha_x]^{W_{xr}(N)}\right) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{x,r\in C} W_{xr}(N) log(a(x\mid r)\alpha_x) \\
&= \sum_{x,r\in C} \lim_{N\to\infty} \frac{W_{xr}(N)}{N} log(a(x\mid r)\alpha_x).
\end{aligned}
$$

Después, por *la ley fuerte de los grandes números* (ver apéndice [Ley fuerte de los grandes números]).

$$
G = \sum_{x,r\in C} p(x,r) log(a(x\mid r)\alpha_x)
$$

casi seguramente. Como $\alpha_x = \frac{1}{p(x)}$, entonces

$$
\begin{aligned}
G &= \sum_{x,r\in C} p(x,r) log\left(\frac{a(x\mid r)}{p(x)}\right) \\
&= \sum_{x,r\in C} p(x,r)\left(log(a(x\mid r)) - log(p(x))\right) \\
&= \sum_{x,r\in C} p(x,r) log(a(x\mid r)) - \sum_{x,r\in C} p(x,r) log(p(x)).
\end{aligned}
$$

Se busca la proporción de capital $a(x \mid r)^*$, para toda $x, r \in C$ que maximice la tasa G. Se toma $r \in C$ fija pues las proporciones de capital $a(x \mid r)$ y $a(x \mid e)$ no guardan ninguna relación cuando $r \neq e$ y se maximiza para cada caracter recibido. Por la proposición dos, el término que maximiza G es

$$\begin{aligned} a(x \mid r) =& \frac{\sum_{x\in C} a(x \mid r)}{\sum_{x\in C} p(x,r)} p(x,r) \\ =& \frac{p(x,r)}{\sum_{x\in C} p(x,r)}, \end{aligned}$$

por propiedades de probabilidad conjunta y la definición de probabilidad condicional

$$a(x \mid r) = \frac{p(x,r)}{q(r)} = q(x \mid r).$$

Entonces

$$G^* = \sum_{x,r\in C} p(x,r) log(q(x \mid r)) - \sum_{x,r\in C} p(x,r) log(p(x)).$$

Considerando X como la variable aleatoria que representa la entrada al canal n-ario de comunicación con ruido, Y como la variable aleatoria que representa la salida de dicho canal y las definiciones de entropía para la variable aleatoria X y la entropía condicional de X dado que se conoce Y, la tasa G^* equivale a la tasa de transmisión de información definida por Shannon

$$\begin{aligned} H(X) =& - \sum_{x,r\in C} p(x,r) log(p(x)) \\ H_Y(X) =& - \sum_{x,r\in C} p(x,r) log(q(x \mid r)) \\ R =& H(X) - H_Y(X) = G^*. \end{aligned}$$

El caso en que las proporciones de pago no son justas, es decir, $\alpha_s \neq \frac{1}{p(s)}$ para alguna $s \in C$, no es muy distinto, la tasa de crecimiento exponencial G puede escribirse como sigue

$$
\begin{aligned}
G &= \sum_{x,r \in C} p(x,r) log(a(x \mid r)\alpha_x) \\
&= \sum_{x,r \in C} p(x,r) log(a(x \mid r)) + \sum_{x,r \in C} p(x,r) log(\alpha_x) \\
&= \sum_{x,r \in C} p(x,r) log(a(x \mid r)) + \sum_{x \in C} p(x) log(\alpha_x)
\end{aligned}
$$

de nuevo, la proporción de capital $a(x \mid r)$ que maximiza G, para cada $r \in C$ es $a^*(x \mid r) = q(x \mid r)$ y

$$
\begin{aligned}
G^* =& \sum_{x,r \in C} p(x,r) log(q(x \mid r)) + \sum_{x \in C} p(x) log(\alpha_x) \\
=& H(\alpha) - H_Y(X)
\end{aligned}
$$

donde

$$
H(\alpha) = \sum_{x \in C} p(x) log(\alpha_x).
$$

El apostador puede omitir si las proporciones de pago son justas o no, la proporción de capital óptima es la misma.

Al igual que en el caso binario se puede interpretar G como una «tasa». Considere i una tasa de interés capitalizable en cada apuesta de tal manera que $i = 2^G - 1$. Sustituyéndose G^*, en el caso de proporciones de pago justas, en la igualdad anterior

$$
\begin{aligned}
i^* =& 2^{\left(\sum_{x,r\in C} p(x,r) log(q(x|r)) - \sum_{x,r\in C} p(x,r) log(p(x))\right)} - 1 \\
=& 2^{\left(\sum_{x,r\in C} log\left(q(x|r)^{p(x,r)}\right) + \sum_{x,r\in C} log\left(p(x)^{-p(x,r)}\right)\right)} - 1 \\
=& 2^{\left(\sum_{x,r\in C} log\left(q(x|r)^{p(x,r)}\right)\right)} 2^{\left(\sum_{x,r\in C} log\left(p(x)^{-p(x,r)}\right)\right)} - 1 \\
=& \prod_{x,r\in C} 2^{log\left(q(x|r)^{p(x,r)}\right)} 2^{log\left(p(x)^{-p(x,r)}\right)} - 1 \\
=& \prod_{x,r\in C} \left(\frac{q(x \mid r)}{p(x)}\right)^{p(x,r)} \\
=& \prod_{x,r\in C} (q(x \mid r)\alpha_x)^{p(x,r)} - 1.
\end{aligned}
$$

Cuando las proporciones de pago no son justas

$$
\begin{aligned}
i^* =& 2^{\left(\sum_{x,r\in C} p(x,r) log(q(x|r)) + \sum_{x,r\in C} p(x,r) log(\alpha_x)\right)} - 1 \\
=& \prod_{x,r\in C} 2^{log\left(q(x|r)^{p(x,r)}\right)} 2^{log\left(\alpha_x^{p(x,r)}\right)} - 1 \\
=& \prod_{x,r\in C} (q(x \mid r)\alpha_x)^{p(x,r)} - 1.
\end{aligned}
$$

El capital crece como una inversión con tasa de interés compuesto $\left(\prod_{x,r\in C} (q(x \mid r)\alpha_x)^{p(x,r)} - 1\right)$ capitalizable en cada apuesta para ambos casos.

Capítulo 3

Criterio de Kelly bayesiano

1. RWIRE

Por una Caminata Aleatoria en un Ambiente Aleatorio (RWIRE, por sus siglas en inglés *Random Walk In a Random Enviroment*) se entiende una pareja $(\{S_n\}_{n=0}^{\infty}, \{\Delta_n\}_{n=0}^{\infty})$ donde $\{S_n\}_{n=0}^{\infty}$ será una caminata aleatoria y $\{\Delta_n\}_{n=0}^{\infty}$ es una familia de variables aleatorias idénticamente distribuídas de tal forma que $0 \leq \Delta_i \leq 1$ y se cumple que $\mathbb{P}[S_{n+1} = i + 1 | S_n = i] = \Delta_i = 1 - \mathbb{P}[S_{n+1} = i - 1 | S_n = i]$.

Coloquialmente se llamará a $\{\Delta_n\}_{n=0}^{\infty}$ el ambiente de la RWIRE y por un ambiente fijo de la RWIRE se entenderá una muestra $\overline{\alpha} = \{\alpha_n\}_{n=0}^{\infty}$ de $\{\Delta_n\}_{n=0}^{\infty}$.

Por lo que para un ambiente fijo se tiene una caminata aleatoria $\{S_n^{\overline{\alpha}}\}_{n=0}^{\infty}$ tal que $\mathbb{P}[S_{n+1} = i + 1 | S_n = i] = \alpha_i$, $1 - \alpha_i = \beta_i = \mathbb{P}[S_{n+1} = i - 1 | S_n = i]$

Considere $\{\alpha_i\}_{i=0}^{\infty}$ una sucesión tal que $\alpha_0 = 1$, $0 < \alpha_i < 1$ para toda $j \geq 1$ y una cadena de Markov con espacio de estados $E = \mathbb{N}$ y matriz de transición como sigue

$$p_{ij} = \begin{cases} 1 & i = 0, j = 1 \\ \alpha_i & j = i + 1 \\ 1 - \alpha_i & j = i - 1 \\ 0 & e.o.c. \end{cases}$$

La cadena descrita es una generalización al problema de la ruina del jugador. La condición $\alpha_0 = 1$ significa que 0 es una barrera reflectiva hacia la derecha. Se podría remplazar a 0 por cualquier entero y ampliar así el esquema de la cadena.

Sea $f_{i,j} = \mathbb{P}[\tau_j < \infty | X_0 = i]$ donde $\tau_i = min\{n \geq 1 | X_n = i\}$ es un tiempo de paro. Entonces f_{ij} representa la probabilidad de llegar a j desde i en un tiempo finito. Por la propiedad fuerte de Markov (ver apéndice [Propiedad fuerte de Markov]), para cualesquiera dos estados $i, j \in E$ con $i < j$ se tiene lo siguiente

$$\begin{aligned}
f_{i,j} =& \mathbb{P}[\tau_j < \infty | X_0 = i] \\
=& \sum_{n=1}^{\infty} \mathbb{P}[\tau_j = n | X_0 = i] \\
=& \sum_{n=1}^{\infty} \sum_{k \in E \setminus \{j\}} \mathbb{P}[\tau_j = n | X_1 = k] p_{ik} \\
=& \sum_{n=1}^{\infty} (\mathbb{P}[\tau_j = n | X_1 = i+1]\alpha_i + \mathbb{P}[\tau_j = n | X_1 = i-1]\beta_i) \\
=& \alpha_i f_{i+1,j} + \beta_i f_{i-1,j},
\end{aligned}$$

además,

$$f_{j,j} = \mathbb{P}[\tau_j < \infty | X_0 = j] = 1, \quad para\, toda\; j \in E.$$

De esta manera, si $j \in E$ entonces se tiene un sistema de ecuaciones con $j + 1$ incognitas

$$\begin{cases} f_{i,j} = \alpha_i f_{i+1,j} + \beta_i f_{i-1,j}, & 0 \leq i < j, \\ f_{j,j} = 1 \end{cases}$$

como $\alpha_0 = 1$, entonces $f_{0,j} = f_{1,j}$. Por otro lado

$$f_{i,j} = (\alpha_i + \beta_i) f_{i,j} = \alpha_i f_{i,j} + \beta_i f_{i,j}$$

entonces, el sistema puede reescribirse como sigue

$$\alpha_i (f_{i+1,j} - f_{i,j}) = \beta_i (f_{i,j} - f_{i-1,j}).$$

Cuya solución es

$$f_{i,j} = 1, \quad 0 \leq i \leq j$$

que es de esperarse ya que 0 es una barrera reflectiva hacia la derecha.

Sean $i, j, k \in E$ tal que $j < k$ y ${}_{jk}f_i = \mathbb{P}[\tau_j < \tau_k | X_0 = i]$. ${}_{jk}f_i$ es la probabilidad de visitar primero el estado j antes que el estado k partiendo desde i. Por la propiedad fuerte de Markov (ver apéndice [Propiedad fuerte de Markov]), si $j < i < k$ entonces

$$\begin{aligned}
{}_{jk}f_i =& \mathbb{P}[\tau_j < \tau_k | X_0 = i] \\
=& \sum_{l \in E \setminus \{j,k\}} \mathbb{P}[\tau_j < \tau_k | X_1 = l] p_{il} \\
=& \mathbb{P}[\tau_j < \tau_k | X_1 = i+1]\alpha_i + \mathbb{P}[\tau_j < \tau_k | X_1 = i-1]\beta_i \\
=& {}_{jk}f_{i+1}\alpha_i + {}_{jk}f_{i-1}\beta_i.
\end{aligned}$$

Además

$${}_{jk}f_j = 1, \quad {}_{jk}f_k = 0.$$

Así, dados $j, k \in E$ con $j < k$ se tiene un sistema de ecuaciones de $k-j+1$ incognitas

$$\begin{cases}
{}_{jk}f_i = {}_{jk}f_{i+1}\alpha_i + {}_{jk}f_{i-1}\beta_i, & j < i < k, \\
{}_{jk}f_j = 1, \\
{}_{jk}f_k = 0.
\end{cases}$$

Como ${}_{jk}f_i = {}_{jk}f_i\alpha_i + {}_{jk}f_i\beta_i$ para $j < i < k$, entonces

$$\begin{aligned}
\alpha_i({}_{jk}f_{i+1} - {}_{jk}f_i) =& \beta_i({}_{jk}f_i - {}_{jk}f_{i-1}) \\
{}_{jk}f_{i+1} - {}_{jk}f_i =& \frac{\beta_i}{\alpha_i}({}_{jk}f_i - {}_{jk}f_{i-1}),
\end{aligned}$$

sustituyendo ${}_{jk}f_i - {}_{jk}f_{i-1}, {}_{jk}f_{i-1} - {}_{jk}f_{i-2}, \ldots, {}_{jk}f_{j+1} - {}_{jk}f_j$ en la expresión anterior se tiene

$$\begin{aligned}
{}_{jk}f_{i+1} - {}_{jk}f_i =& \frac{\beta_i\beta_{i-1}\ldots\beta_{j+1}}{\alpha_i\alpha_{i-1}\ldots\alpha_{j+1}}({}_{jk}f_{j+1} - {}_{jk}f_j) \\
{}_{jk}f_{i+1} - {}_{jk}f_i =& \frac{\beta_i\beta_{i-1}\ldots\beta_{j+1}}{\alpha_i\alpha_{i-1}\ldots\alpha_{j+1}}({}_{jk}f_{j+1} - 1) \quad (*),
\end{aligned}$$

para $j < i \leq k$. Haciendo

$$\gamma_j = 1, \quad \gamma_i = \frac{\beta_i ... \beta_{j+1}}{\alpha_i ... \alpha_{j+1}}.$$

Sumando las diferencias de (*)

$$\begin{aligned}\sum_{r=j}^{i-1} ({}_{jk}f_{r+1} - {}_{jk}f_r) &= \sum_{r=j}^{i-1} \gamma_r ({}_{jk}f_{j+1} - 1)\\ {}_{jk}f_i - 1 &= \sum_{r=j}^{i-1} \gamma_r ({}_{jk}f_{j+1} - 1), \quad j < i \leq k \quad (**),\end{aligned}$$

en particular, para k se tiene

$$\begin{aligned}{}_{jk}f_k - 1 &= \sum_{r=j}^{k-1} \gamma_r ({}_{jk}f_{j+1} - 1)\\ -1 &= \sum_{r=j}^{k-1} \gamma_r ({}_{jk}f_{j+1} - 1)\\ \frac{-1}{\sum_{r=j}^{k-1} \gamma_r} &= {}_{jk}f_{j+1} - 1.\end{aligned}$$

Sustituyendo la expresión anterior en (**) se tiene

$$\begin{aligned}{}_{jk}f_i - 1 =& \frac{-\sum_{r=j}^{i-1} \gamma_r}{\sum_{r=j}^{k-1} \gamma_r}\\ {}_{jk}f_i =& 1 - \frac{\sum_{r=j}^{i-1} \gamma_r}{\sum_{r=j}^{k-1} \gamma_r}\\ {}_{jk}f_i =& \frac{\sum_{r=i}^{k-1} \gamma_r}{\sum_{r=j}^{k-1} \gamma_r}, \quad j < i < k,\end{aligned}$$

además, $\mathbb{P}[\tau_j < \tau_k | X_0 = i] = 1 - \mathbb{P}[\tau_k < \tau_j | X_0 = i]$, entonces

$$_{kj}f_i = \frac{\sum_{r=j}^{i-1} \gamma_r}{\sum_{r=j}^{k-1} \gamma_r}, \quad j < i < k.$$

2. Teorema de De Finetti

DEFINICIÓN 1. *Una sucesión de variables aleatorias* $\{X_n\}_{n=1}^{\infty}$ *es intercambiable si para toda* $n \geq 2$

$$X_1, \ldots, X_n \stackrel{D}{=} X_{\pi(1)}, \ldots, X_{\pi(n)},$$

para toda $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ *permutación.*

PROPOSICIÓN 3. *Una sucesión de variables aleatorias independientes e idénticamente distribuídas,* $\{X_n\}_{n=1}^{\infty}$*, es intercambiable.*
Dem. Sean $n \geq 2$*,* $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ *permutación y* F *la función de distribución común de* X_i*, entonces*

$$\begin{aligned}\mathbb{P}[X_1 \leq x_1, \ldots, X_n \leq x_n] =& \prod_{i=1}^{n} F(x_i) \\ =& \prod_{i=1}^{n} F(x_{\pi(i)}) \\ =& \mathbb{P}[X_1 \leq x_{\pi(1)}, \ldots, X_n \leq x_{\pi(n)}].\end{aligned}$$

□

El recíproco no es cierto de manera general como se muestra en el siguiente ejemplo.

EJEMPLO 1. *(Urna de Polya). Considere una urna con* b *bolas blancas y* n *bolas negras. Se toma un bola de manera aleatoria, se nota su color y se remplaza con la bola y* $a > 0$ *bolas del mismo color. Sea* $X_i = 1$ *si la i-ésima bola que se saca es negra y* $X_i = 0$ *en otro caso. Se afirma que la sucesión* $\{X_i\}_{i=1}^{\infty}$ *es intercambiable.*
Sean $m \geq 2$*,* $k = \sum_{i=1}^{m} x_i$ *y* $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ *permutación, entonces*

$$\mathbb{P}[x_1, \ldots, x_m] = \frac{\prod_{i=0}^{k-1}(n+ia) \prod_{j=0}^{m-k-1}(b+ja)}{\prod_{l=0}^{m-1}(n+b+la)} = \mathbb{P}[x_{\pi(1)}, \ldots, x_{\pi(m)}].$$

Por otro lado, basta probar que X_1 *no es independiente a* X_2 *para demostrar que la sucesión no es independendiente.*

$$\mathbb{P}[X_1 = 1, X_2 = 1] = \tfrac{n}{n+b} \tfrac{n+a}{n+b+a},$$

y

$$\begin{aligned}\mathbb{P}[X_1=1]\mathbb{P}[X_2=1]=&\frac{n}{n+b}\left(\mathbb{P}[X_1=0]\mathbb{P}[X_2=1|X_1=0]+\mathbb{P}[X_1=1]\mathbb{P}[X_2=1|X_1=1]\right)\\=&\frac{n}{n+b}\left(\frac{b}{n+b}\frac{n}{n+b+a}+\frac{n}{n+b}\frac{n+a}{n+b+a}\right)\\\neq&\mathbb{P}[X_1=1,X_2=1].\end{aligned}$$

La sucesión de variables aleatorias de la urna de Polya no es independiente y no son un proceso de Markov.

Teorema 1. *(De Finetti). Una sucesión binaria de variables aleatorias* $\{X_n\}_{n=1}^{\infty}$ *es intercambiable si y solo si existe una función de distribución* F *en el* $[0,1]$ *tal que para toda* n

$$\mathbb{P}[X_1=x_1,\dots,X_n=x_n]=\int_0^1\theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}dF(\theta).$$

3. Caminata aleatoria

Se considera el caso en el que $\Delta_n=\alpha$ para toda $n\geq 1$, es decir, se «fija» un ambiente constante . Sean $h:[0,1]\to\mathbb{R}$ la función de densidad *a priori* de α, $\{Z_j\}_{j=1}^{\infty}$ una sucesión de variables aleatorias tales que $\mathbb{P}[Z_i=1|\alpha]=\alpha=1-\mathbb{P}[Z_i=-1|\alpha]$ y considere $\{S_n\}_{n=0}^{\infty}$ una sucesión de variables aleatorias tal que $S_0=0$, $S_n=\sum_{i=1}^n Z_i$, $W_i=(Z_i+1)/2$ y $Y_n=(S_n+n)/2=\sum_{i=1}^n W_i$. Note que $W_i|\alpha\sim Blli(\alpha)$, entonces, si se supone independencia de la sucesión $\{W_i|\alpha\}_{i=1}^{\infty}$ se tiene que $Y_n|\alpha\sim Bin(n,\alpha)$. Sea $y=\sum_{i=1}^n w_i$, entonces se tiene lo siguiente (ver apéndice [Estadística bayesiana]).

$$\begin{aligned}f(w_1,\dots,w_n,\alpha)=&\prod_{i=1}^n f(w_i|\alpha)h(\alpha)\\=&\prod_{i=1}^n\alpha^{w_i}(1-\alpha)^{1-w_i}h(\alpha)\\=&\alpha^y(1-\alpha)^y h(\alpha).\end{aligned}$$

Como $\alpha\in[0,1]$, entonces

$$f(x_1,\dots,x_n)=\int_0^1\alpha^y(1-\alpha)^{n-y}h(\alpha)d\alpha.$$

De esta manera

$$f(\alpha|Y_n)=\frac{\alpha^{Y_n}(1-\alpha)^{n-Y_n}h(\alpha)}{\int_0^1\alpha^{Y_n}(1-\alpha)^{n-Y_n}h(\alpha)d\alpha}.$$

Con momentos

$$\mathbb{E}[\alpha^j|Y_n] = \frac{\int_0^1 \alpha^{Y_n+j}(1-\alpha)^{n-Y_n}h(\alpha)d\alpha}{\int_0^1 \alpha^{Y_n}(1-\alpha)^{n-Y_n}h(\alpha)d\alpha}.$$

Esta distribución es única para una sucesión de variables aleatorias intercambiables por el teorema de De Finetti.

PROPOSICIÓN 4. *$\{S_n\}$ es una cadena de Markov con matriz de transición $\mathbb{P}[S_{n+1} = s_n + 1|S_n = s_n] = \mathbb{E}[\alpha|S_n = s_n]$.*
Primero se encuentra la probabilidad de transición $\mathbb{P}[S_{n+1} = s_n + 1|S_n = s_n]$

$$\begin{aligned}
\mathbb{P}[S_{n+1} = s_n + 1|S_n = s_n] =& \frac{\mathbb{P}[Z_{n+1} = 1, S_n = \sum_{i=1}^n z_i]}{\mathbb{P}[S_n = \sum_{i=1}^n z_i]} \\
=& \frac{\int_0^1 f(Z_{n+1} = 1|\alpha) \prod_{i=1}^n f(z_i|\alpha)h(\alpha)d\alpha}{\int_0^1 \prod_{i=1}^n f(z_i|\alpha)h(\alpha)d\alpha} \\
=& \frac{\int_0^1 \alpha^{(s_n+n)/2+1}(1-\alpha)^{(n-s_n)/2}h(\alpha)d\alpha}{\int_0^1 \alpha^{(s_n+n)/2}(1-\alpha)^{(n-s_n)/2}h(\alpha)d\alpha} \\
=& \mathbb{E}[\alpha|S_n = s_n],
\end{aligned}$$

por otro lado

$$\begin{aligned}
\mathbb{P}[S_{n+1}|S_n = s_n, \ldots, S_0 = s_0] =& \frac{\mathbb{P}[S_{n+1}, s_n, \ldots, s_0]}{\mathbb{P}[s_n, \ldots, s_0]} \\
=& \frac{\mathbb{P}[\sum_{i=1}^{n+1} Z_i, s_n, \ldots, S_0]}{\mathbb{P}[s_n, \ldots, s_0]} \\
=& \frac{\mathbb{P}[Z_{n+1}, z_n, \ldots, z_0]}{\mathbb{P}[z_n, \ldots, z_0]} \\
=& \mathbb{P}[S_{n+1}|S_n = s_n].
\end{aligned}$$

Por lo tanto $\{S_n\}_{n=0}^\infty$ es un proceso de Markov. □

PROPOSICIÓN 5. *Cuando la distribución a priori es Beta la distribución a posteriori vuelve a ser Beta.*
Suponga que $\alpha \sim Beta(a, b)$, i.e.

$$h(\alpha) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\alpha^{a-1}(1-\alpha)^{b-1}; \quad a, b > 0, \quad \alpha \in (0, 1),$$

con media y varianza

$$\mathbb{E}[\alpha] = \frac{a}{a+b}, \quad Var(\alpha) = \frac{ab}{(a+b)^2(a+b+1)},$$

respectivamente.
La función de distribución a posteriori, $\alpha|Y_n$ está dada por

$$
\begin{aligned}
f(\alpha|Y_n) =& \frac{\alpha^{Y_n}(1-\alpha)^{n-Y_n}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\alpha^{a-1}(1-\alpha)^{b-1}}{\int_0^1 \alpha^{Y_n}(1-\alpha)^{n-Y_n}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\alpha^{a-1}(1-\alpha)^{b-1}d\alpha} \\
=& \frac{\alpha^{Y_n+a-1}(1-\alpha)^{n+b-Y_n-1}}{\int_0^1 \alpha^{Y_n+a-1}(1-\alpha)^{n+b-Y_n-1}d\alpha} \\
=& \frac{\frac{\Gamma(n+a+b)}{\Gamma(a+Y_n)\Gamma(n+b-Y_n)}\alpha^{Y_n+a-1}(1-\alpha)^{n+b-Y_n-1}}{\int_0^1 \frac{\Gamma(n+a+b)}{\Gamma(a+Y_n)\Gamma(n+b-Y_n)}\alpha^{Y_n+a-1}(1-\alpha)^{n+b-Y_n-1}d\alpha} \\
=& \frac{\Gamma(n+a+b)}{\Gamma(a+Y_n)\Gamma(n+b-Y_n)}\alpha^{Y_n+a-1}(1-\alpha)^{n+b-Y_n-1},
\end{aligned}
$$

i.e. $\alpha|Y_n \sim Beta(Y_n+a, n+b-Y_n)$, *con media y varianza posteriores*

$$\mathbb{E}[\alpha|Y_n] = \tfrac{Y_n+a}{n+b+a}, \quad Var(\alpha|Y_n) = \tfrac{(Y_n+a)(n+b-Y_n)}{(n+a+b)^2(n+a+b+1)}.$$

□

Se reescribe la esperanza condicional de $\alpha|Y_n$ como sigue

$$
\begin{aligned}
\mathbb{E}[\alpha|Y_n] =& \frac{a}{n+a+b} + \frac{Y_n}{n+b+a} \\
=& \left(\frac{a+b}{n+a+b}\right)\left(\frac{a}{a+b}\right) + \left(\frac{n}{n+b+a}\right)\left(\frac{Y_n}{n}\right) \\
=& \frac{a+b}{n+b+a}\mathbb{E}[\alpha] + \frac{n}{n+b+a}\left(\frac{Y_n}{n}\right).
\end{aligned}
$$

Que es una combinación convexa de la media *a priori* y la media observada. Por la ley de los grandes números (ver apéndice [Ley fuerte de los grandes números]) $Y_n/n \to \alpha$ con probabilidad uno cuando $n \to \infty$, entonces

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[\alpha|Y_n] =& \lim_{n\to\infty}\left(\frac{a+b}{n+b+a}\mathbb{E}[\alpha] + \frac{n}{n+b+a}\left(\frac{Y_n}{n}\right)\right) \\
=& \lim_{n\to\infty}\frac{Y_n}{n} \\
=& \alpha,
\end{aligned}
$$

casi seguramente.

Como $Z_i = 2W_i - 1$ y $S_n = \sum_{i=1}^{n} Z_i = 2Y_n - n$, entonces

$$\begin{aligned}\mathbb{P}[S_{n+1} = S_n + 1|S_n] =&\mathbb{E}[\alpha|S_n]\\ =&\frac{\frac{S_n+n}{2} + a}{n+b+a}\\ =&\frac{S_n+n+2a}{2(n+b+a)}.\end{aligned}$$

Además, la esperanza y varianza de S_n están dadas por

$$\mathbb{E}[S_n] = \mathbb{E}[\mathbb{E}[S_n|\alpha]] = n(2\mathbb{E}[\alpha] - 1) = n\left(\tfrac{2a}{a+b} - 1\right) = n\tfrac{a-b}{a+b},$$

$$\begin{aligned}Var(S_n) =&\mathbb{E}[Var(S_n|\alpha)] + Var(\mathbb{E}[S_n|\alpha])\\ =&4n\mathbb{E}[\alpha(1-\alpha)] + 4n^2Var(\alpha)\\ =&4n\left(\mathbb{E}^[\alpha] - \mathbb{E}[\alpha^2]\right) + 4n^2Var(\alpha)\\ =&4n\left(\frac{a}{a+b} - \frac{a(a^2+a(b+1)+b)}{(a+b+1)(a+b)^2} + n\frac{ab}{(a+b+1)(a+b)^2}\right)\\ =&4n\frac{a(a+b)(a+b+1) - ab - a^2(a+b+1) + nab}{(a+b+1)(a+b)^2}\\ =&4n\frac{a^3+a^2+2a^2b+ab^2+ab-ab-a^3-a^2b-a^2+nab}{(a+b+1)(a+b)^2}\\ =&\frac{4nab(a+b+n)}{(a+b+1)(a+b)^2}.\end{aligned}$$

4. Apuestas

Sean K_0 el capital inicial del apostador, K_N el capital después de N apuestas, f_i la fracción de capital apostada en la i-ésima apuesta y $(\{S_n\}_{n=0}^{\infty}, \{\Delta_n\}_{n=0}^{\infty})$ una RWIRE como en la sección anterior. Note que el capital del apostador a tiempo j es $K_j = K_{j-1} + K_{j-1}f_jZ_j = K_{j-1}(1 + f_jZ_j)$, y de manera recursiva se tiene que el capital del apostador a tiempo N es

$$K_N = K_0 \prod_{i=1}^{N}(1 + f_iZ_i).$$

El objetivo es maximizar $\mathbb{E}[ln(K_N)]$ sobre todas las fracciones de capital $f_1, \ldots, f_N$.

Considere $\{\mathcal{F}_i\}_{i=0}^{\infty}$ una filtración tal que $\mathcal{F}_i = \sigma(\{Z_0, \ldots, Z_i\})$. Note que f_n debe ser adaptado a $\mathcal{F}_{n-1}$, es decir, f_n depende a lo más de la información actual y entonces de los valores observados $Z_0, \ldots, Z_{n-1}$. Primero se considerará una distribución *a priori* general. Así las cosas, se tiene

$$\begin{aligned}
\underset{f_1,\ldots,f_N}{max}\ \mathbb{E}[ln(K_N)] =& \underset{f_1,\ldots,f_N}{max}\ \mathbb{E}\left[ln\left(K_0\prod_{i=1}^{N}(1+f_iZ_i)\right)\right]\\
=& ln\,K_0 + \sum_{i=1}^{N}\underset{f_i}{max}\,\mathbb{E}[ln(1+f_iZ_i)]\\
=& ln\,K_0 + \sum_{i=1}^{N}\underset{f_i}{max}\,\mathbb{E}[\mathbb{E}[ln(1+f_iZ_i)|\mathcal{F}_{i-1}]]\\
=& ln\,K_0 + \mathbb{E}\left[\sum_{i=1}^{N}\underset{f_i}{max}\,\mathbb{E}[ln(1+f_iZ_i)|\mathcal{F}_{i-1}]\right].
\end{aligned}$$

Que en el caso markoviano se reduce a

$$\underset{f_1,\ldots,f_N}{max}\ \mathbb{E}[ln(K_N)] = ln\ K_0 + \mathbb{E}[\textstyle\sum_{i=1}^{N}\underset{f_i}{max}\,\mathbb{E}[ln(1+f_iZ_i)|S_{i-1}]]. \quad (1)$$

La probabilidad de transición de la RWIRE se denotará como $\mathbb{P}[S_{k+1} = S_k + 1|S_k] = \mathbb{E}[\alpha|S_k, k] = 1 - \mathbb{P}[S_{k+1} = S_k - 1|S_k]$.

Para continuar se supone que se han observado k apuestas y considere

$$F_n(x, S_k, k) = \underset{f_{k+1},\cdots,f_{k+n}}{max}\ \mathbb{E}[ln(K_{k+n})|K_k = x, S_k, k].$$

Procediendo de manera análoga a (1) se tiene lo siguiente

$$F_n(x, S_k, k) = ln(x) + \mathbb{E}[\sum_{i=k}^{k+n-1}\underset{f_i}{max}\,\mathbb{E}[ln(1+f_{i+1}Z_{i+1})|S_i, i]].$$

Después, para cada i se tiene

$$\begin{aligned}
\mathbb{E}[ln(1+f_{i+1}Z_{i+1})|S_i, i] =& \mathbb{P}[Z_{i+1} = 1|S_i, i]ln(1+f_{i+1}) + \mathbb{P}[Z_{i+1} = -1|S_i, i]ln(1-f_{i+1})\\
=& \mathbb{E}[\alpha|S_i, i]ln(1+f_{i+1}) + (1-\mathbb{E}[\alpha|S_i, i])ln(1-f_{i+1}). \quad (2)
\end{aligned}$$

Haciendo uso de la Proposición dos , el máximo valor de $\mathbb{E}[ln(1+f_{i+1}Z_{i+1})|S_i,i]$ se alcanza en x^* dado por

$$x^* = (2\mathbb{E}[\alpha|S_i,i], 2(1-\mathbb{E}[\alpha|S_i,i])).$$

Es decir,

$$\begin{cases} 1+f_{i+1} = 2\mathbb{E}[\alpha|S_i,i], \\ 1-f_{i+1} = 2(1-\mathbb{E}[\alpha|S_i,i]), \end{cases}$$

cuya solución es

$$f_{i+1}^* = 2\mathbb{E}[\alpha|S_i,i]-1,$$

sustituyendo f_{i+1}^* en (2) se obtiene

$$\begin{aligned}\mathbb{E}[ln(1+f_{i+1}Z_{i+1})|S_i,i] =& \mathbb{E}[\alpha|S_i,i]ln(1+2\mathbb{E}[\alpha|S_i,i]-1)+(1-\mathbb{E}[\alpha|S_i,i])ln(1-2\mathbb{E}[\alpha|S_i,i]+1) \\ =& \mathbb{E}[\alpha|S_i,i]ln(2\mathbb{E}[\alpha|S_i,i])+(1-\mathbb{E}[\alpha|S_i,i])ln(2(1-\mathbb{E}[\alpha|S_i,i])) \\ =& ln(2)+\mathbb{E}[\alpha|S_i,i]ln(\mathbb{E}[\alpha|S_i,i])+(1-\mathbb{E}[\alpha|S_i,i])ln(1-\mathbb{E}[\alpha|S_i,i]).\end{aligned}$$

En el caso en que la distribución *a priori* es beta(a,b) lo anterior se reduce a lo siguiente

$$\begin{aligned} f_{i+1}^* =& 2\left(\frac{S_i+i+2a}{2(i+a+b)}\right)-1 \\ =& \frac{S_i+a-b}{i+a+b}, \end{aligned}$$

y

$$\begin{aligned} &\mathbb{E}[ln(1+f_{i+1}Z_{i+1})|S_i,i] \\ &= \frac{S_i+i+2a}{2(i+a+b)}ln\left(\frac{S_i+i+2a}{i+a+b}\right)+\left(1-\frac{S_i+i+2a}{2(i+a+b)}\right)ln\left(2-\frac{S_i+i+2a}{i+a+b}\right) \\ &= \frac{S_i+i+2a}{2(i+a+b)}ln\left(\frac{S_i+i+2a}{i+a+b}\right)+\frac{i-S_i+2b}{2(i+a+b)}ln\left(\frac{i-S_i+2b}{i+a+b}\right). \end{aligned}$$

5. El criterio

El siguiente teorema es valido suponiendo que apostar en contra está permitido en un canal de comunicación binario [**B**, pp. 65-78].

TEOREMA 2. *Suponga una serie de N apuestas donde la probabilidad de ganar la i-ésima apuesta es p_i y sea K_N el capital del apostador después de N apuestas, entonces $\mathbb{E}[ln(K_N)]$ se maximiza al apostar en el i-ésimo juego la proporción de capital óptima de acuerdo al Criterio de Kelly: $l_i = 2p_i - 1$.*

El teorema anterior permite realizar apuestas en series de juegos más complejos como BlackJack o deportes donde las probabilidades de victoria pueden cambiar de un juego a otro.

De esta forma, suponiendo una RWIRE como en este capítulo y el teorema anterior se obtiene el siguiente resultado.

TEOREMA 3. *Suponga que se han observado j resultados de una RWIRE con ambiente «fijo» constante α en una serie de N apuestas y el capital de un apostador es x. Entonces, para $i = j, \ldots, N-1$, la estrategia óptima consiste en apostar la proporción del Criterio de Kelly correspondiente a la información determinada antes del $i+1$-ésimo juego, es decir $f_{i+1} = 2\mathbb{E}[\alpha|S_i, i] - 1$. Además, en la $i+1$-ésima apuesta el valor esperado del logaritmo del capital es*

$$ln(2) + \mathbb{E}[\alpha|S_i, i]ln(\mathbb{E}[\alpha|S_i, i]) + (1 - \mathbb{E}[\alpha|S_i, i])ln(1 - \mathbb{E}[\alpha|S_i, i]).$$

Finalmente, suponiendo $\alpha \sim Beta(a, b)$ entonces

$$f_{i+1} = \frac{S_i + a - b}{i + a + b},$$

y el valor esperado del logaritmo del capital en la $i+1$-ésima apuesta es

$$\frac{S_i + i + 2a}{2(i + a + b)} ln\left(\frac{S_i + i + 2a}{i + a + b}\right) + \frac{i - S_i + 2b}{2(i + a + b)} ln\left(\frac{i - S_i + 2b}{i + a + b}\right).$$

□

Capítulo 4

Simulación Monte Carlo

Para probar el modelo se simuló una caminata aleatoria con probabilidad de victoria $p = 0.55, 0.65, 0.75, 0.85, 0.95$ y, con base en la caminata, el logaritmo natural del capital de J usando el criterio de Kelly bayesiano suponiendo una distribución *a priori* beta (ver apéndice [Código]). Apostar en contra está permitido.

Suponiendo que no hay observaciones *a priori* el capital inicial apostado es $f = \frac{1}{2}$. Además, suponiendo que se conoce p, se simuló el logaritmo natural del capital de J usando el criterio de Kelly.

Considere las Figuras 2-6 donde se grafica el logaritmo natural del capital de J usando el criterio de Kelly bayesiano contra el criterio de Kelly no bayesiano en una serie de 2000 apuestas.

FIGURA 2. Logaritmo natural del capital de J usando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano cuando $p = 0.55$

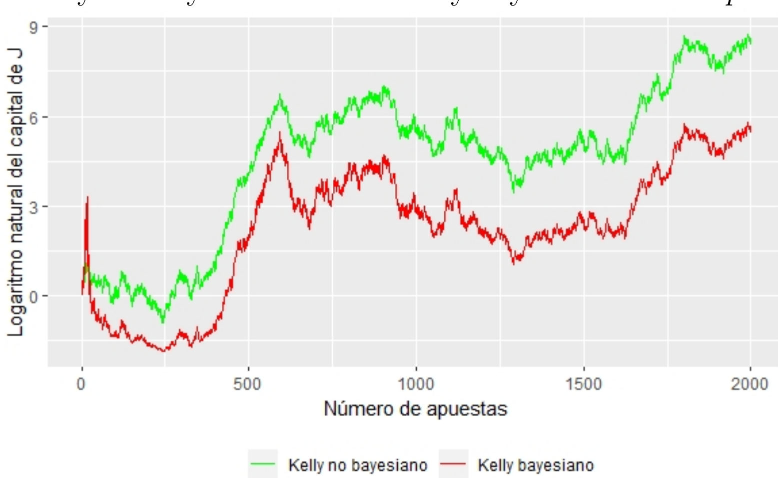


FIGURA 3. Logaritmo natural del capital de J usando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano cuando $p = 0.65$

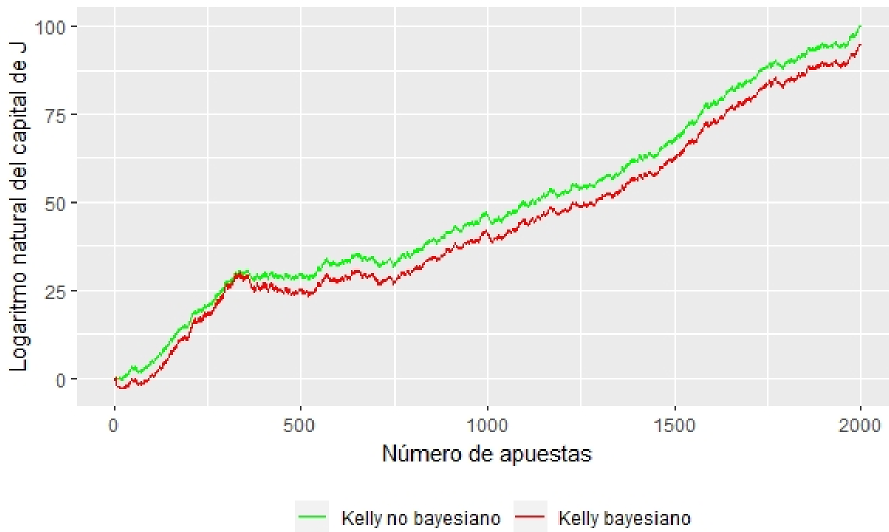


FIGURA 4. Logaritmo natural del capital de J usando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano cuando $p = 0.75$

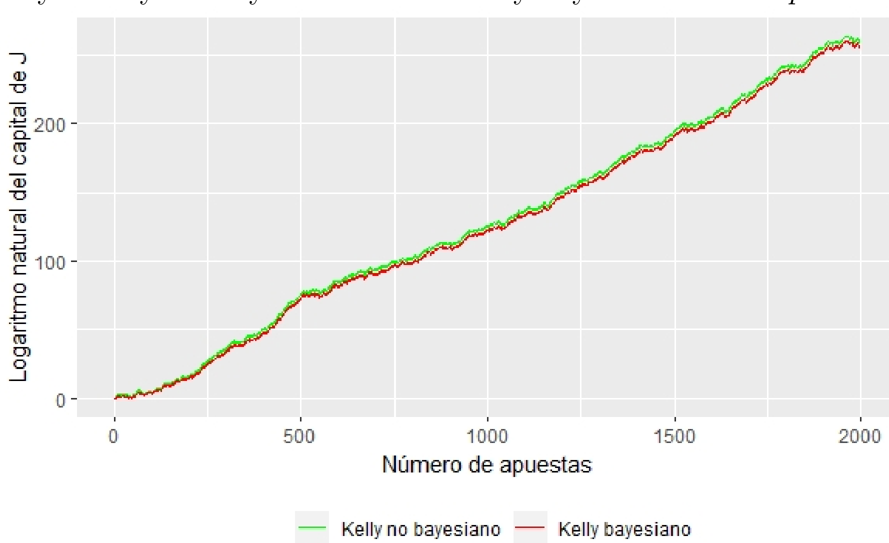


FIGURA 5. Logaritmo natural del capital de J usando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano cuando $p = 0.85$

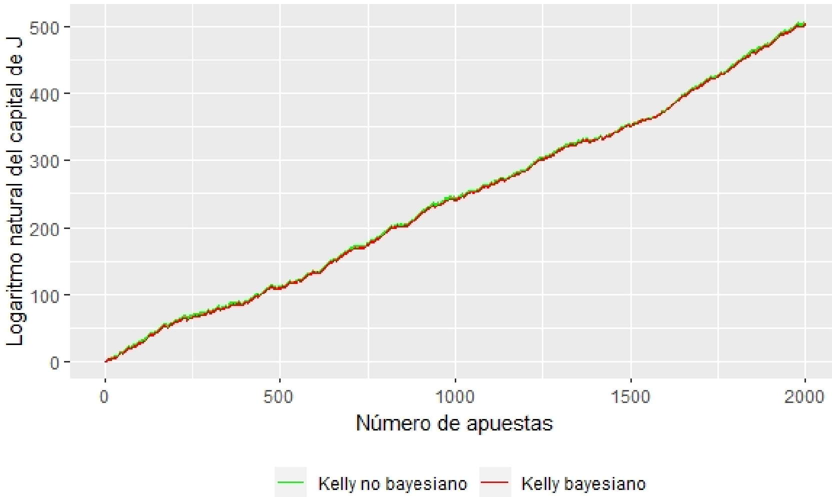
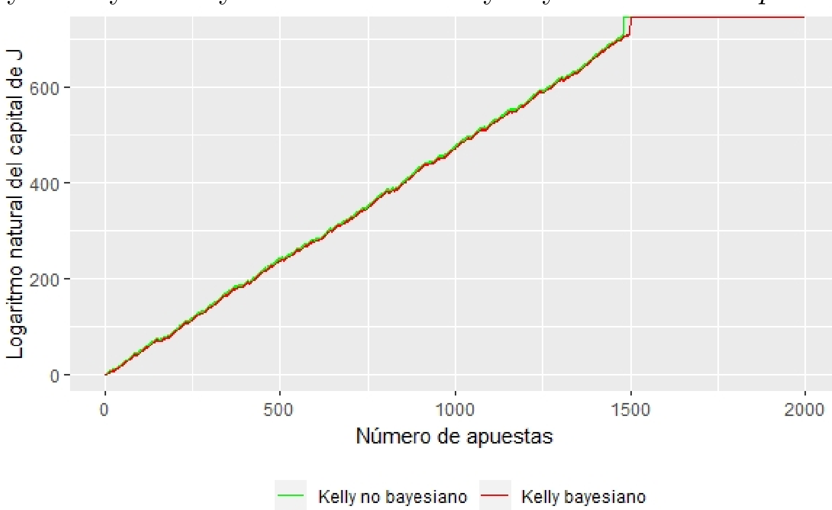


FIGURA 6. Logaritmo natural del capital de J usando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano cuando $p = 0.95$



Note que la probabilidad de victoria es estimada más «rápido» cuando es mayor y en consecuencia la utilidad utilizando Kelly bayesiano difiere menos de la utilidad utilizando el criterio Kelly no bayesiano.

Finalmente se realizó una simulación Monte Carlo de tamaño 1000 en la que se calculó el MAPE (*Mean Absolute Percentage Error*) entre el logaritmo natural del capital de J usando el criterio Kelly no bayesiano contra el criterio de Kelly bayesiano cuando $p = 0.55, 0.65, 0.75, 0.85, 0.95$ (ver apéndice [Código]).

El MAPE se define como sigue

$$M = \frac{100\,\%}{n} \sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right|.$$

Donde A_i es el valor real observado del i-ésimo experimento, F_i el valor estimado del i-ésimo experimento y n el número de experimentos realizados.

Si K_n y $\overline{K_n}$ denotan los capitales después de n apuestas de J utilizando el criterio de Kelly no bayesiano y el criterio de Kelly no bayesiano, respectivamente, el MAPE, M_p, que se simuló para cada p se calculó de la siguiente manera

$$M_p = \frac{100\,\%}{n} \sum_{i=1}^{n} \left| \frac{ln(K_i) - ln(\overline{K_i})}{ln(K_i)} \right|.$$

En cada simulación se tomo como base una trayectoria de 500 pasos de una caminata aleatoria, es decir $n = 500$.

Considere la tabla uno donde se muestra el promedio de los M_p.

$\overline{M_{0.55}}$	$\overline{M_{0.65}}$	$\overline{M_{0.75}}$	$\overline{M_{0.85}}$	$\overline{M_{0.95}}$
1055.31 %	86.48 %	23.28 %	10.56 %	6.36 %

TABLA 1. *Mean Absolute Percentge Error*

De nuevo se observa que mientras la probabilidad de victoria es mayor la diferencia entre las log-utilidades del capital de J utilizando el criterio de Kelly no bayesiano y el criterio de Kelly bayesiano es menor, con un rango de aproximadamente de 1049 % entre los promedios de los M_p con $p = 0.55$ y $p = 0.95$.

Capítulo 5

Aplicaciones

1. Deportes

Con base en el histórico de las temporadas regulares de equipos de la *National Football League*, *Major League of Baseball* y *National Basketball Association* se simuló el capital de un apostador J con tres estrategias distintas suponiendo que apostar en contra está permitido.

Para la *NFL* se tomaron los resultados de la temporada 2018-2019 de los *New England Patriots*, *Los Angeles Rams*, *Kansas City Chiefs* y *New Orleans Saints* para simular el capital de J.

Para la *MLB* se tomaron los resultados de la temporada 2018 de los *Boston Red Sox*, *New York Yankees*, *Tampa Bay Rays* y *Chicago Cubs* para simular el capital de J.

Para la *NBA* se tomaron los resultados de a temporada 2018-19 de los *Golden State Warriors*, *Milwaukee Buck*, *Los Angeles Lakers* y *Toronto Raptors* para simular el capital de J

Para la primera estrategia, *criterio de Kelly bayesiano*, se supuso una disribución *a priori* beta con parámetros a y b iguales al número de victorias y derrotas, respectivamente, de cada equipo durante la temporada regular anterior. Por ejemplo, los parámetros que se consideraron para la distribución *a priori* de los *Boston Red Sox* fueron $a = 93$ y $b = 69$. Para simular el capital de J con esta estrategia se supuso $K_0 = 1$ y para $n \geq 1$ se tiene $K_n = K_{n-1}(1+f_n)$ si el equipo ganó o $K_n = K_{n-1}(1-f_n)$ si el equipo perdió, donde f_n es la proporción de capital del criterio Kelly bayesiano después de $n-1$ apuestas.

La segunda estrategia, *media a priori*, consistió en apostar la proporción del criterio de Kelly suponiendo que la probabilidad de victoria de cada equipo era igual al número de victorias de la temporada regular anterior dividido por el total de juegos de la temporada. Por ejemplo, para los *Golden State Warriors* se consideró $p = \frac{67}{82}$. Para simular el capital de J con esta estrategia se supuso $K_0 = 1$ y para $n \geq 1$ se tiene $K_n = K_{n-1}(1+f)$ si el equipo ganó o $K_n = K_{n-1}(1-f)$ si el equipo perdió, donde f es la proporción de capital del criterio Kelly no bayesiano asociada a p.

Finalmente, para la tercera estrategia, *criterio de Kelly no bayesiano*, se conoce el número de victorias al final de la temporada en curso y se apostó la proporción del criterio de Kelly asociada a la probabilidad de victoria. Por ejemplo, para los *New England Patriots* se consideró $p = \frac{11}{16}$. Para simular el capital de J con esta estrategia se supuso $K_0 = 1$ y para $n \geq 1$ se tiene $K_n = K_{n-1}(1+f)$ si el equipo ganó o $K_n = K_{n-1}(1-f)$ si el equipo perdió, donde f es la proporción de capital del criterio Kelly no bayesiano asociada a p.

Considere las figuras 7-10

Figura 7. Capital de J apostando a *New England Patriots*

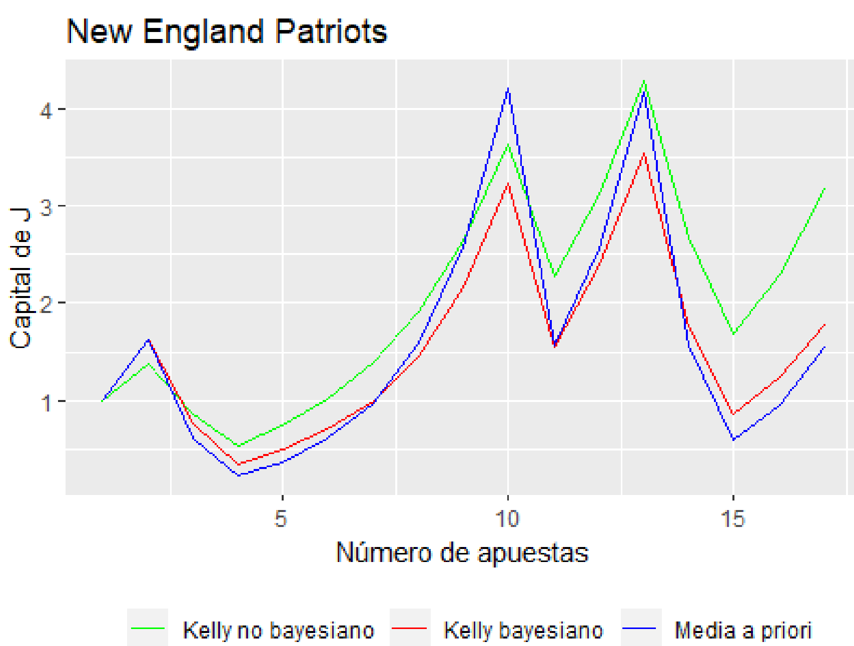


Figura 8. Capital de J apostando a *Los Angeles Rams*

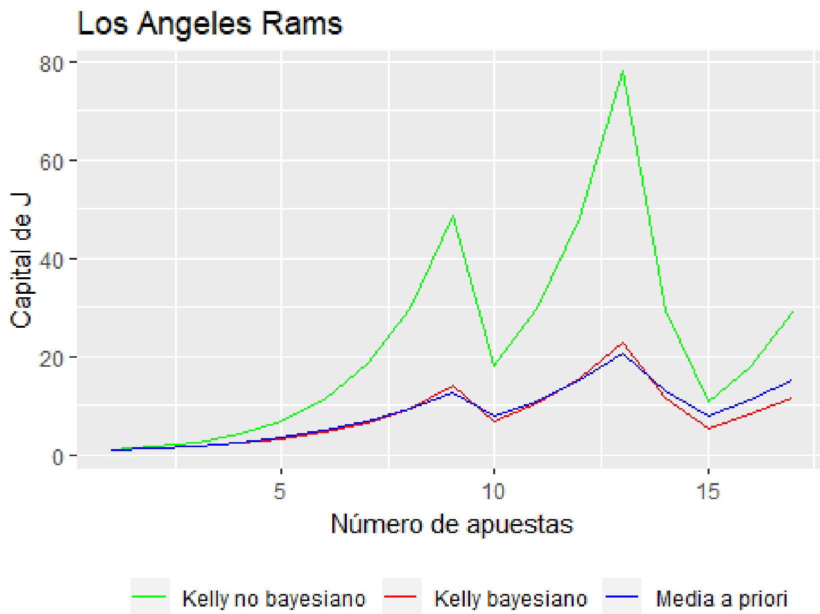


Figura 9. Capital de J apostando a *Kansas City Chiefs*

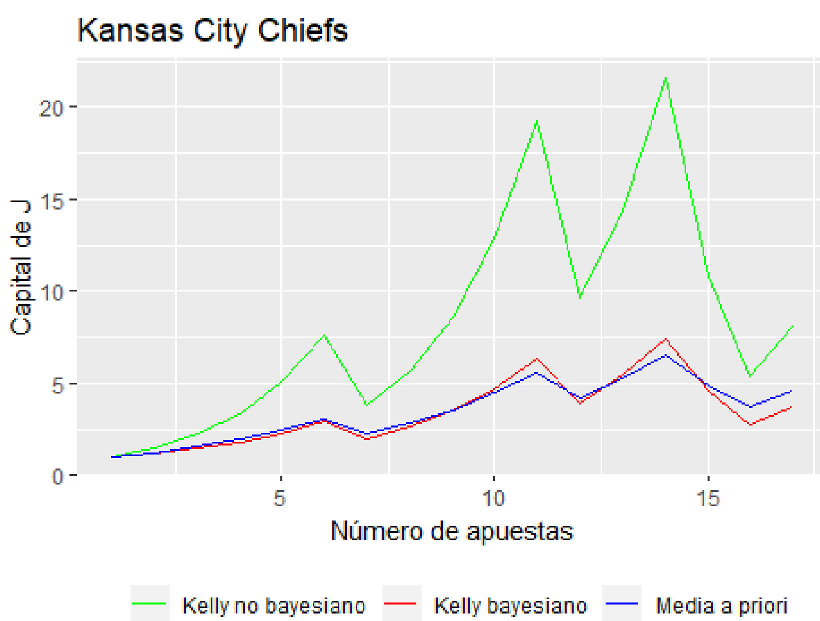
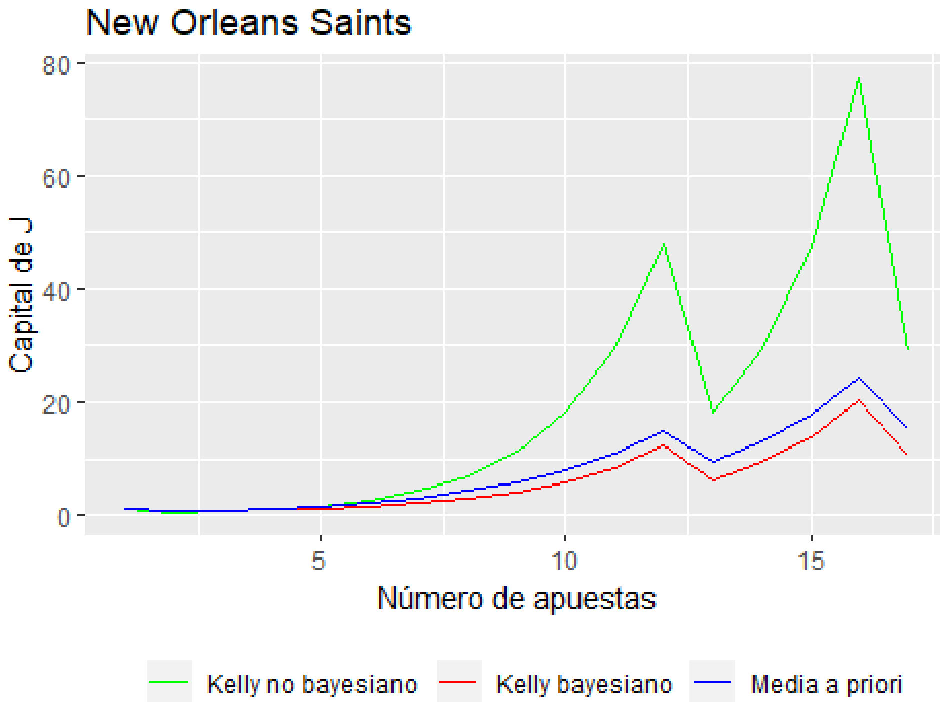


Figura 10. Capital de J apostando a *New Orleans Saints*



Se observa que las tres estrategias resultan en ganancias para J apostando a los cuatro equipos. La estrategia que prevalece para los cuatro equipos es *criterio de Kelly no bayesiano*. Por otro lado, las estrategias *criterio de Kelly bayesiano* y *media a priori* se asemejan a lo largo de la temporada. En la tabla dos se muestra el capital de J al final de los 16 partidos de temporada regular.

Equipo	*Criterio de Kelly bayesiano*	*Media a priori*	*Criterio de Kelly no bayesiano*
Chiefs	3.7638	4.6043	8.1091
Patriots	1.7761	1.5472	3.1676
Rams	11.5448	15.3312	29.0528
Saints	10.4400	15.3312	29.0528

Tabla 2. Capital de J al final de la temporada regular de la *NFL*

Considere las figuras 11-14

FIGURA 11. Capital de J apostando a *Boston Red Sox*

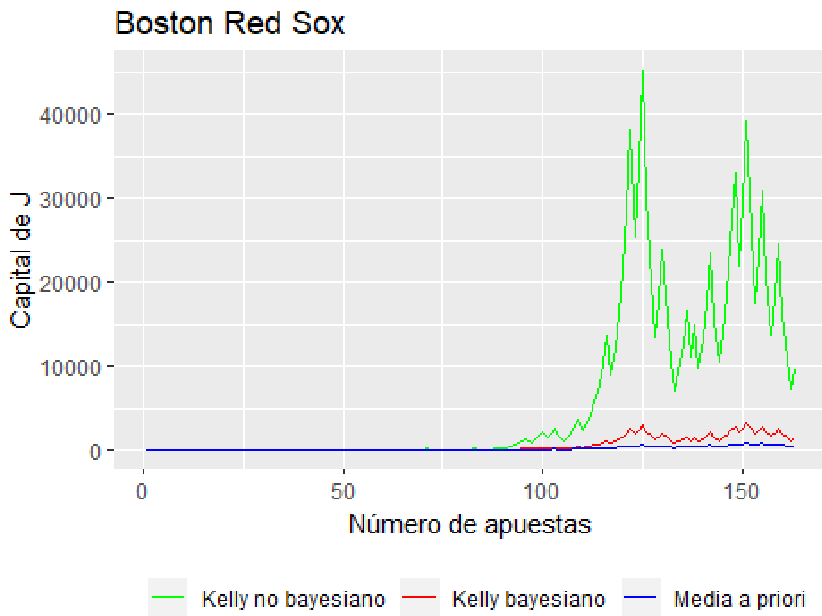


FIGURA 12. Capital de J apostando a *New York Yankees*

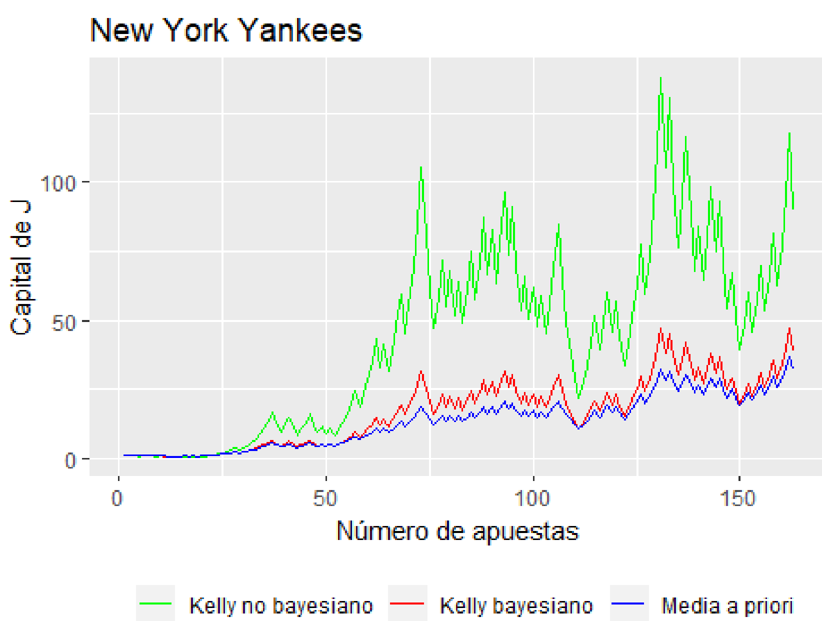


Figura 13. Capital de J apostando a *Tampa Bay Rays*

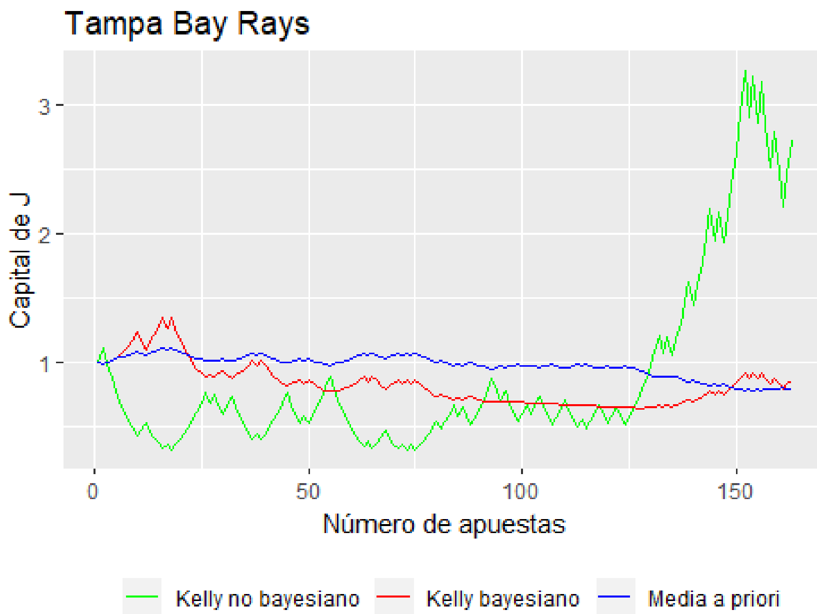
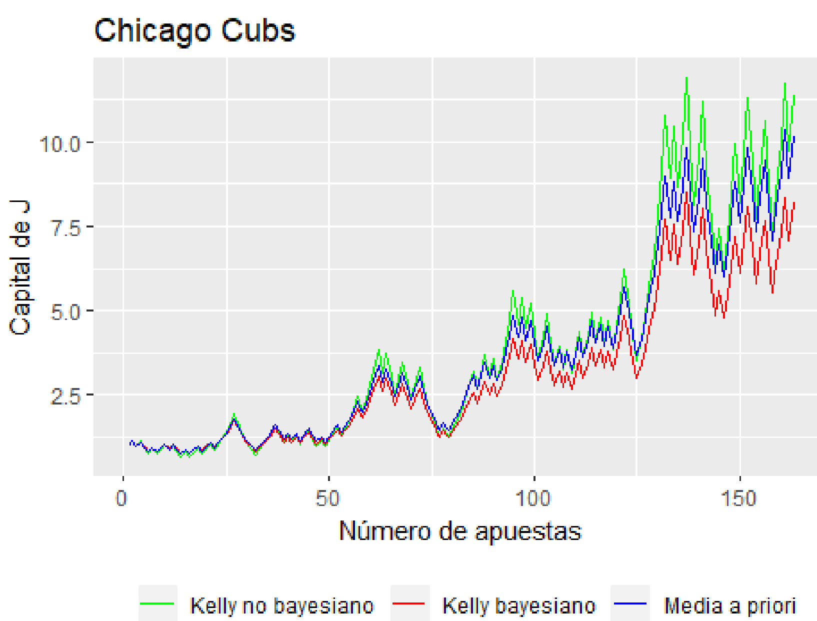


Figura 14. Capital de J apostando a *Chicago Cubs*

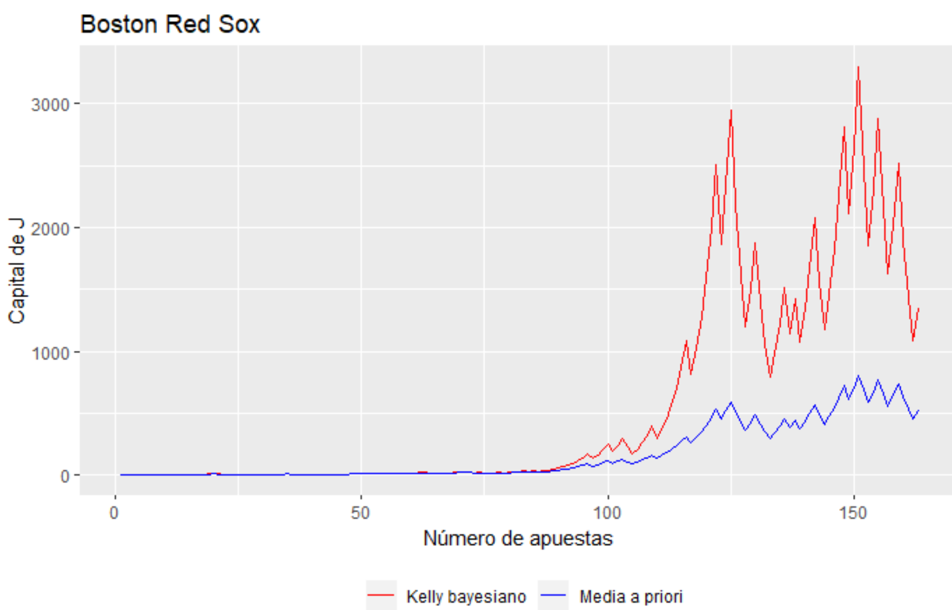


La única estrategia que resulta en ganancias para J apostando a los cuatro equipos es *criterio de Kelly no bayesiano*, misma estrategia es la que da mayores ganancias a J. Por otro lado, implementar *criterio de Kelly bayesiano* y *media a priori* resulta en pérdidas cuando J apuesta a *Tampa Bay Rays* y en ganancias para el resto de los equipos. En la tabla tres se muestra el capital de J al final de los 162 partidos de temporada regular. Adicionalmente, en la figura 15 se comparan las estrategias *criterio de Kelly bayesiano* y *media a priori* cuando J apuesta a *Boston Red Sox*.

Equipo	*Criterio de Kelly bayesiano*	*Media a priori*	*Criterio de Kelly no bayesiano*
Cubs	8.1771	10.1587	11.3810
Rays	0.8651	0.7909	2.7239
Red Sox	1344.3603	524.1267	9648.06
Yankees	38.9204	32.1798	89.8901

TABLA 3. Capital de J al final de la temporada regular de la *MLB*

FIGURA 15. Capital de J apostando a *Boston Red Sox* utilizando las estrategias *criterio de Kelly bayesiano* y *media a priori*



Considere las figuras 16-19

FIGURA 16. Capital de J apostando a *Golden State Warriors*

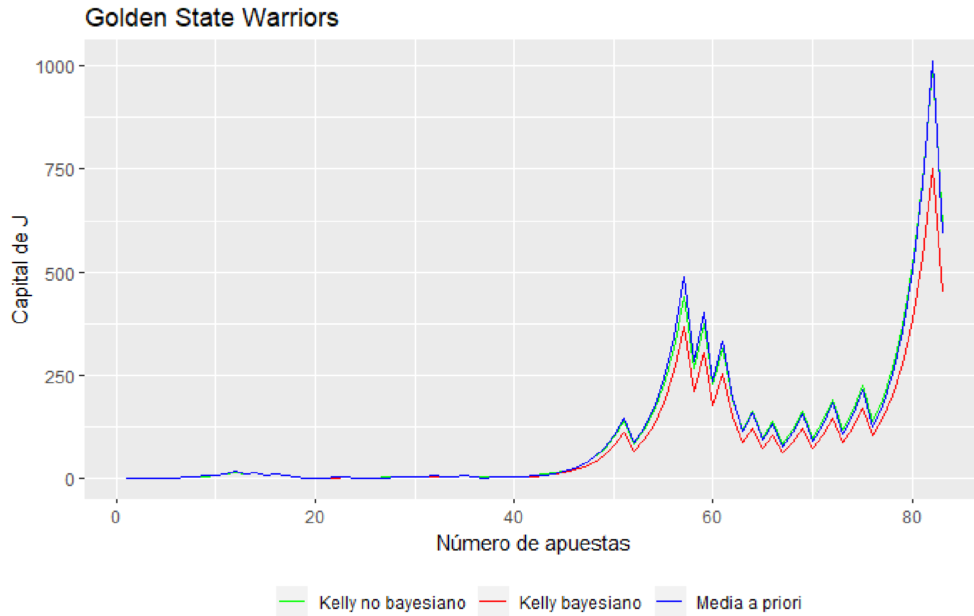


FIGURA 17. Capital de J apostando a *Milwaukee Bucks*

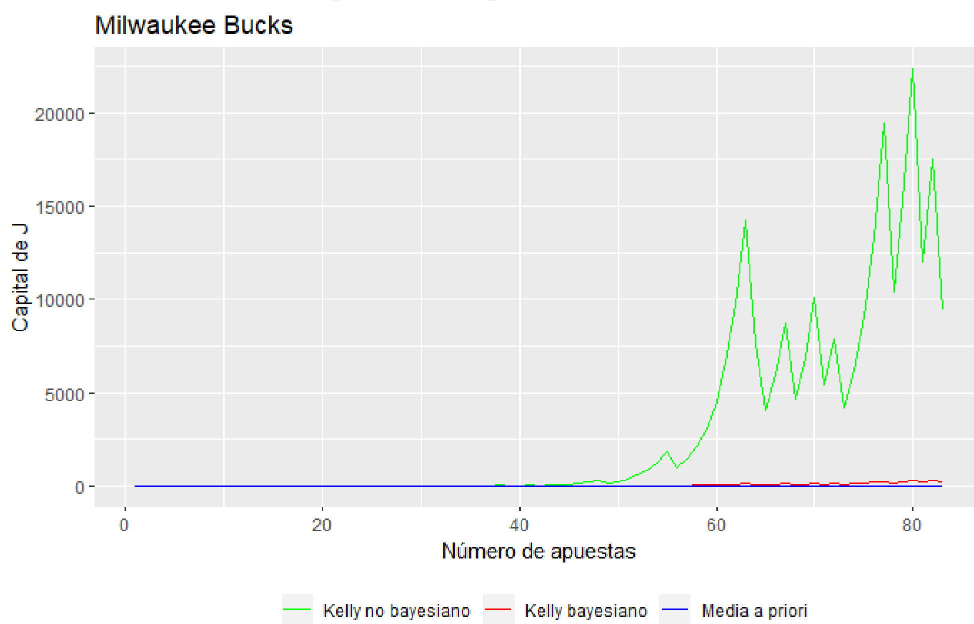


FIGURA 18. Capital de J apostando a *Los Angeles Lakers*

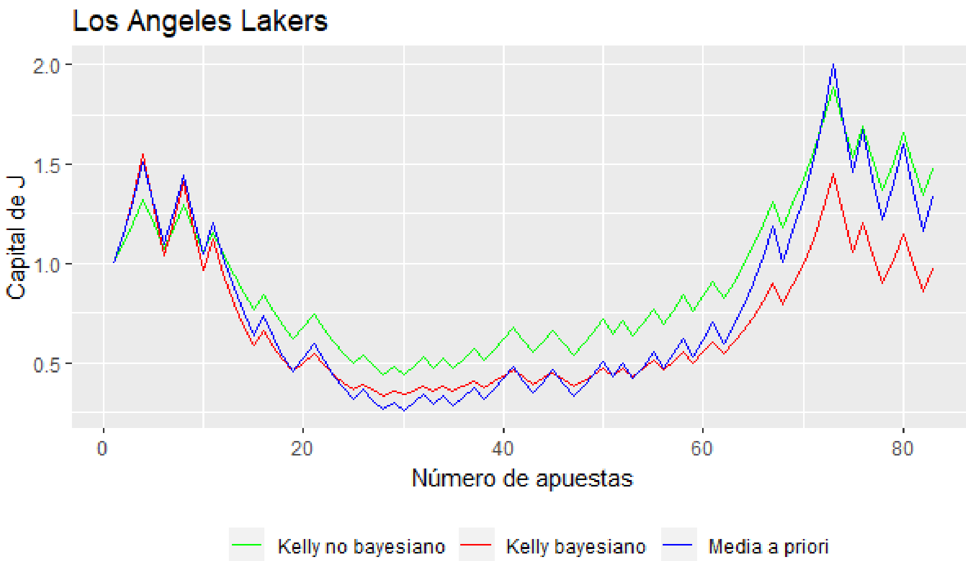
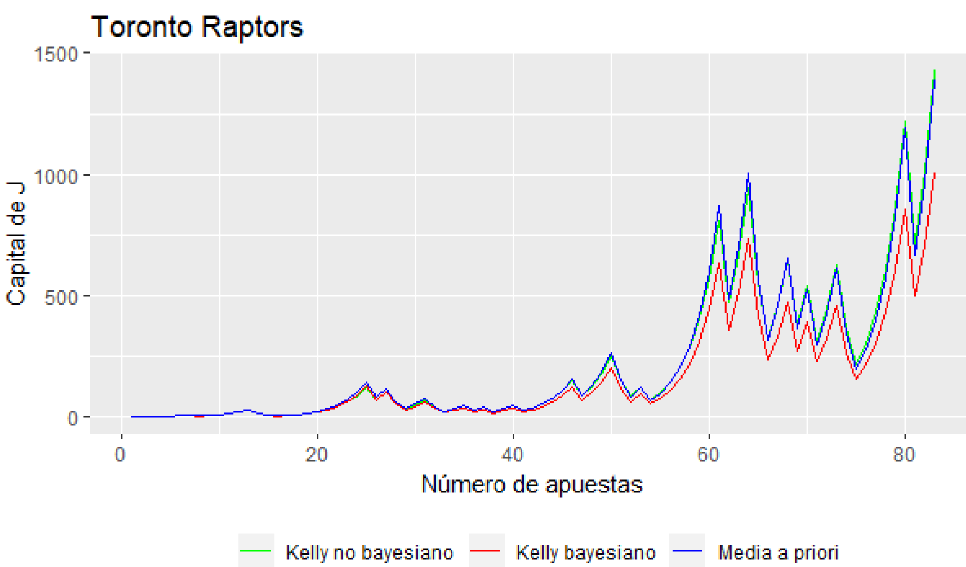


FIGURA 19. Capital de J apostando a *Toronto Raptors*



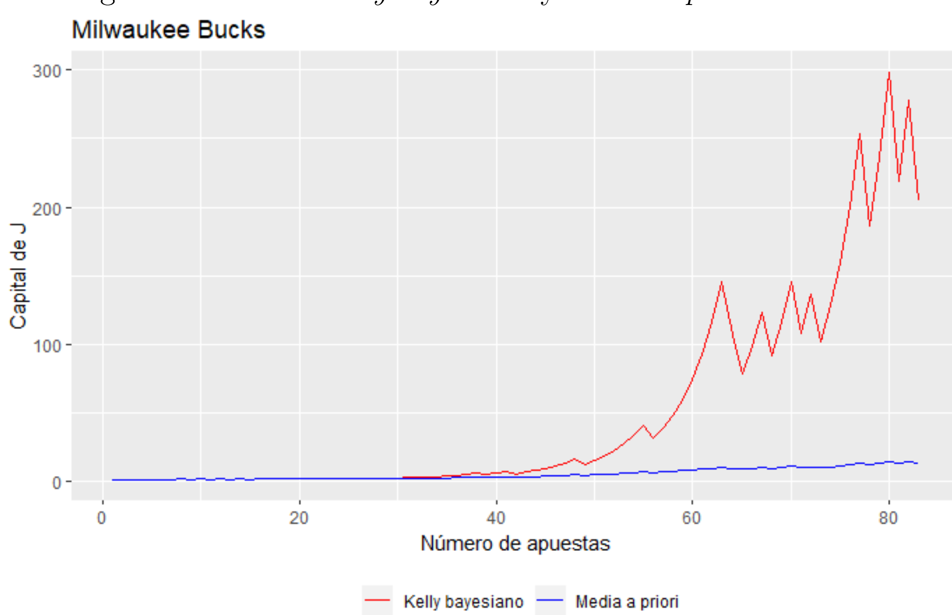
Las estrategias *criterio de Kelly no bayesiano* y *media a priori* resultan en ganancias para J apostando a los cuatro equipos. Al utilizar *criteio de Kelly bayesiano*, J

tiene pérdidas al apostar a *Los Angeles Lakers* y ganancias al apostar al resto de los equipos. En la tabla cuatro se muestra el capital de J al final de los 82 partidos de temporada regular. Adicionalmente, en la figura 20 se comparan las estrategias *criteio de Kelly bayesiano* y *media a priori* cuando J apuesta a *Milwaukee Bucks*.

Equipo	*Criterio de Kelly bayesiano*	*Media a priori*	*Criterio de Kelly no bayesiano*
Bucks	204.8120	13.0054	9421.0162
Lakers	0.9729	1.3386	1.4783
Raptors	1002.2617	1388.7002	1430.9260
Warriors	451.1248	592.1073	609.6643

TABLA 4. Capital de J al final de la temporada regular de la *NBA*

FIGURA 20. Capital de J apostando a *Milwaukee Bucks* utilizando las estrategias *criterio de Kelly bayesiano* y *media a priori*



Capítulo 6

Conclusiones

El criterio de Kelly es una estrategia teórica elegante y eficaz para cualquier apostador o inversionista, sin embargo, como se remarcó antes, conocer la probabilidad de victoria es un supuesto muy fuerte en la práctica. Así, un apostador, que busca maximizar el valor esperado de su log-utilidad no puede utilizar de forma directa el criterio de Kelly y debe buscar alternativas.

El criterio, como se demostró en el capítulo tres, otorga una estrategia tan elegante como el criterio de Kelly para maximizar la log-utilidad mientras estima la probabilidad de victoria. Además, la simulación Monte Carlo realizada en el cuarto capítulo comprueba la eficiencia del criterio bayesiano, obteniendo mejores resultados cuando la probabilidad de victoria es mayor.

El capítulo cinco aterriza toda la teoría desarrollada y los resultados son positivos de manera general. La estrategia adaptativa que ofrece el criterio de Kelly bayesiano resulta en ganancias muy por encima de estrategias fijas en ligas con temporadas regulares largas.

De cómo Edward Thorp implementó el criterio de Kelly para obtener las enormes ganancias para él y sus inversionistas se sabe poco. A saber, es igual de probable que conociera las probabilidades victoria en cada apuesta que realizaba a que tuviera la mejor de las suertes. Con los resultados obtenidos en este trabajo introductorio espero que el lector este convencido de que Thorp se pudo haber ayudado del criterio de Kelly bayesiano para tener una carrera tan redituable como la de pocos en los deportes, BlackJack y la bolsa de valores.

Apéndice

1. Ley fuerte de los grandes números

Teorema 4. *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes e identicamente distribuidas con media finita μ. Entonces*

$$\lim_{n\to\infty} \tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i \stackrel{c.s.}{\to} \mu$$

Dem. Para cada n se tiene $X_n = X_n^+ - X_n^-$ donde X_n^+ y X_n^- representan las partes positiva y negativa de X_n, respectivamente. Como $\mathbb{E}[X_1] < \infty$ entonces $\mathbb{E}[X_n^+], \mathbb{E}[X_n^+] < \infty$ para toda n. Después,

$$\tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i = \tfrac{1}{n}\sum_{i=1}^{n} X_i^+ - \tfrac{1}{n}\sum_{i=1}^{n} X_i^-$$

y en consecuencia, basta demostrar el teorema para variables aleatorias positivas.

Considere $Y_n = X_n\mathbf{1}_{\{X_n\leq n\}}$ para cada n, entonces

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{P}[X_n \neq Y_n] =& \sum_{n=1}^{\infty}\mathbb{P}[X_n > Y_n] = \sum_{n=1}^{\infty}\mathbb{P}[X_1 > Y_n]\\ \leq& \sum_{n=1}^{\infty}\int_{n-1}^{n}\mathbb{P}[X_1 > t]dt\\ =& \int_0^{\infty}\mathbb{P}[X_1 > t]dt\\ =&\mathbb{E}[X_1] < \infty.\end{aligned}$$

*Por el lema de Borel-Cantelli [**R4**] se tiene que, con probabilidad 1, $X_n = Y_n$ para toda n excepto una cantidad finita de ellas.*
En consecuencia

$$\tfrac{1}{n}\textstyle\sum_{i=1}^{n}(Y_i - X_i) \to 0 \quad c.p.\,1$$

Además

$$\mathbb{E}[Y_n] = \mathbb{E}[X_n\mathbf{1}_{\{X_n\leq n\}}] = \mathbb{E}[X_1\mathbf{1}_{\{X_n\leq n\}}] \to \mathbb{E}[X_1].$$

Entonces

$$\mathbb{E}[\bar{Y}_n] = \tfrac{1}{n}\textstyle\sum_{i=0}^{n}\mathbb{E}[Y_n] \underset{n\to\infty}{\to} \mathbb{E}[X_1]$$

Considere ahora $\alpha > 1$*, la sucesión* $u_k = [\alpha^k]$ *Y* $\epsilon > 0$*. Por la desigualdad de Chebyshev y la independencia de las variables aleatorias* Y_n *se tiene lo siguiente*

$$\begin{aligned}
\mathbb{P}[|\bar{Y}_{u_k} - \mathbb{E}[\bar{Y}_{u_k}]| > \epsilon] &\leq \frac{1}{\epsilon^2} Var(\bar{Y}_{u_k}) \\
&= \frac{1}{\epsilon^2 u_k^2}\sum_{i=1}^{u_k} Var(Y_i) \\
&\leq \frac{1}{\epsilon^2 u_k^2}\sum_{i=1}^{u_k}\mathbb{E}[Y_i^2]
\end{aligned}$$

después,

$$\begin{aligned}
\sum_{k=1}^{\infty}\mathbb{P}[|\bar{Y}_{u_k} - \mathbb{E}[\bar{Y}_{u_k}]| > \epsilon] &\leq \frac{1}{\epsilon^2}\sum_{k=1}^{\infty}\frac{1}{u_k^2}\sum_{i=1}^{u_k}\mathbb{E}[Y_i^2] \\
&= \frac{1}{\epsilon^2}\sum_{i=1}^{\infty}\mathbb{E}[Y_i^2]\sum_{k:u_k\geq i}\frac{1}{u_k^2}
\end{aligned}$$

Como $u_k = [\alpha^k] > \alpha^{k-1}$ *para* $1 < \alpha < \infty$*,*

$$\sum_{k:u_k\geq i}\frac{1}{u_k^2} \leq \sum_{k:\alpha^{k-1}\geq i}\frac{1}{\alpha^{(k-1)2}} \leq \frac{C_1}{i^2}.$$

Para alguna C_1, $0 < C_1 < \infty$. Entonces

$$\begin{aligned}\sum_{k=1}^{\infty} \mathbb{P}[|\bar{Y}_{u_k} - \mathbb{E}[\bar{Y}_{u_k}]| > \epsilon] &\leq \frac{C_1}{\epsilon^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \mathbb{E}[Y_i^2] \\ &= \frac{C_1}{\epsilon^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \mathbb{E}[X_1^2 \mathbf{1}_{\{0 \leq X_1 \leq i\}}] \\ &= \frac{C_1}{\epsilon^2} \sum_{i=1}^{\infty} \sum_{j=1}^{i} \frac{1}{i^2} \mathbb{E}[X_1^2 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}}] \\ &= \frac{C_1}{\epsilon^2} \sum_{j=1}^{\infty} \mathbb{E}[X_1^2 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}}] \sum_{i=j}^{\infty} \frac{1}{i^2} \\ &\leq \frac{C_1}{\epsilon^2} \sum_{j=1}^{\infty} \mathbb{E}[j X_1 \mathbf{1}_{\{j-1 \leq X_1 \leq j\}}] C_2 j^{-1} \\ &= \frac{C_1 C_2}{\epsilon^2} \mathbb{E}[X_1] < \infty.\end{aligned}$$

Esta serie es convergente para toda $\epsilon > 0$. Por el lema de Borel-Cantelli se tiene que

$$\bar{Y}_{u_k} - \mathbb{E}[\bar{Y}_{u_k}] \to 0 \quad c.p.\,1 \quad \textit{cuando } k \to \infty$$

Más aún, lo anterior significa que

$$\bar{Y}_{u_k} \stackrel{c.s.}{\to} \mathbb{E}[X_1] \quad \textit{cuando } k \to \infty$$

Por otro lado, para toda n existe k tal que $u_k \leq n < u_{k+1}$, y como las variables aleatorias Y_n son no negativas se tiene lo siguiente

$$\frac{1}{n} \sum_{i=1}^{u_k} Y_i \leq \bar{Y}_n \leq \frac{1}{n} \sum_{i=1}^{u_{k+1}} Y_i$$

después

$$\frac{u_k}{n} \bar{Y}_{u_k} \leq \bar{Y}_n \leq \frac{u_{k+1}}{n} \bar{Y}_{u_{k+1}}$$

de donde se obtiene que

$$\frac{1}{\alpha} \bar{Y}_{u_k} \leq \bar{Y}_n \leq \alpha \bar{Y}_{u_{k+1}}$$

Haciendo $n \to \infty$, con probabilidad 1 lo siguiente es válido

$$\frac{1}{\alpha} \mathbb{E}[X_1] \leq \liminf_{n \to \infty} \bar{Y}_n \leq \limsup_{n \to \infty} \bar{Y}_n \leq \alpha \mathbb{E}[X_1]$$

Finalmente, haciendo $\alpha \to 1$

$$\bar{Y}_n \stackrel{c.s.}{\to} \mathbb{E}[X_1]$$

□

2. Estadística bayesiana

Sí se supone que α asume una función de densidad $h(\alpha)$, llamada *a priori*, entonces la distribución conjunta de una muestra aleatoria y α, asumiendo independencia de la muestra condicionada a α, está dada por

$$f(x_1,\ldots,x_n,\alpha) = f(x_1,\ldots,x_n|\alpha)h(\alpha) = \prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)$$

La distribución marginal de la muestra es

$$f(x_1,\ldots,x_n) = \int_\alpha \prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)d\alpha$$

Aplicando la fórmula de Bayes

$$\begin{aligned} f(\alpha|x_1,\ldots,x_n) =& \frac{f(x_1,\ldots,x_n|\alpha)h(\alpha)}{f(x_1,\ldots,x_n)} \\ =& \frac{\prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)}{\int_\alpha \prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)d\alpha} \end{aligned}$$

A esta función se le conoce como *a posteriori*. Los momentos están dados por

$$\mathbb{E}[\alpha^n|x_1,\ldots,x_n] = \frac{\int_\alpha \alpha^n \prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)}{\int_\alpha \prod_{i=1}^{n} f(x_i|\alpha)h(\alpha)d\alpha}$$

3. Propiedad fuerte de Markov

En procesos estocásticos es de interés el tiempo que transcurre antes de que ocurra cierto evento.

Definición 2. *Sea $\{X_n\}_{n\in\mathbb{N}}$ un proceso estocástico. Una variable aleatoria $\tau : \Omega \to [0, 1\ldots,\infty]$ es un tiempo de paro respecto a $\{X_n\}_{n\in\mathbb{N}}$ si $(\tau = n) \in \sigma(X_0, X_1,\ldots,X_n)$ para cada $n \in \mathbb{N}$.*

Intuitivamente un tiempo de paro es el instante en que se cumple una condición resultado de una trayectoria del proceso estocástico.

Teorema 5. *(Propiedad fuerte de Markov) Sea $\{X_n\}_{n=0}^{\infty}$ una cadena de Markov y τ un tiempo de paro respecto a este proceso. Condicionado a $(\tau < \infty)$ entonces*

$$\mathbb{P}[X_{\tau+n+1} = j|X_{\tau+n} = i, X_{\tau+n-1} = x_{n-1},\ldots,X_0 = x_0] = \mathbb{P}[X_{\tau+n+1} = j|X_{\tau+n} = i]$$

es decir, $\{X_{\tau+n}\}_{n=0}^{\infty}$ es una cadena de Markov

4. Filtraciones

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Si $\mathcal{F}_1$ y $\mathcal{F}_2$ son dos sub σ-álgebras de $\mathcal{F}$ tales que $\mathcal{F}_1 \subset \mathcal{F}_2$ entonces $\mathcal{F}_2$ tiene más información que $\mathcal{F}_1$ en el sentido de que una σ-álgebra tiene más eventos que otra.

Definición 3. *Una sucesión de σ-álgebras $\{\mathcal{F}_n\}_{n=1}^{\infty}$ es una filtración si $\mathcal{F}_n \subset \mathcal{F}_m \subset \mathcal{F}$ cuando $n < m$.*

Si se interpreta a n como el tiempo entonces $\mathcal{F}_n$ contiene la información disponible a tiempo n.

Definición 4. *Un proceso estocástico $\{X_n\}_{n=0}^{\infty}$ es adaptado a una filtración si $X_n \in \mathcal{F}_n$ para toda $n \geq 1$*

5. Código

Para simular el capital de J usando el criterio de Kelly bayesiano en una serie de $n-1$ apuestas con probabilidad de victoria p se utilizó el siguiente código en R.

```
P=rep.int(0,n)
for(i in 1:n){
        #Experimentos Blli(p)
        P[j]=rbinom(1,1,p)
}

S=rep.int(0,n)
fb=rep.int(1/2,n)
#Parametros Beta
a, b

for(j in 2:n){
        #Caminata aleatoria
        S[j]=S[j-1]+P[j-1]

        #Proporciones de capital Kelly Bayesiano
        fb[j]=(2*S[j]-(j-1)+a-b)/((j-1)+a+b)
}

Kb=rep.int(1,n)

for(j in 2:n){
        #Capital de J
```

```
        Kb[j]=Kb[j-1]*(1+fb[j-1])^(P[j-1])*(1-fb[j-1])^(1-P[j-1])
}
```

Para realizar m simulaciones del MAPE entre el logaritmo natural del capital de J utilizando el criterio de Kelly no bayesiano contra el criterio de Kelly bayesiano en una serie de $n-1$ apuestas con probabilidad de victoria p se utilizó el siguiente código en R.

```
P=rep.int(0,n)
#Parametros Beta
a,b

fb=rep.int(1/2,n)
#Proporcion Kelly No Bayesiano
f=2*p-1
S=rep.int(0,n)
K=rep.int(1,n)
Kb=rep.int(1,n)
MAPE=rep.int(0,m)

for(i in 1:m){
 M=0
 for(j in 1:n){
  #Experimentos Blli(p)
  P[j]=rbinom(1,1,p)

  #Caminata aleatoria
  S[j]=ifelse(j==1,S[1],S[j-1]+P[j-1])

  #Proporciones Kelly Bayesiano
  fb[j]=ifelse(j==1,fb[1],(2*S[k,j]-(j-1)+a-b)/((j-1)+a+b))

  #Capital Kelly No Bayesiano
  K[j]=ifelse(j==1,1,K[j-1]*(1+f[j])^(P[j-1])*(1-f[j])^(1-P[j-1])

  #Capital Kelly Bayesiano
  Kb[j]=ifelse(j==1,1,Kb[j-1]*(1+fb[j-1])^(P[j-1])*
  (1-fb[j-1])^(1-P[j-1]))

  M=ifelse(j==1,0,M+abs(1-log(K)[j]/log(Kb)[j]))
```

```
    }
    MAPE[i]=M/(n-1)
}
```

Para simular el capital de J usando el criterio de Kelly bayesiano, el criterio de Kelly no bayesiano y media a priori durante la temporada regular de un equipo cualquiera se utilizó el siguiente código en R. Para ejemplificar se escribe el código utilizado para *New England Patriots*.

```
#New England Patriots
#Parametros Beta
a=13; b=3

#Media a priori y proporcion de Kelly
m=13/16; fp=2*m-1

#Resultados de la temporada regular
r=c(1,0,0,1,1,1,1,1,1,0,1,1,0,0,1,1)

#Probabilidad de victoria despues de la temporada regular
#y proporcion de Kelly
p=sum(r)/length(r)
f=2*p-1

S=rep.int(0,length(r)+1)
fb=rep.int(fp,length(r)+1)

for(j in 2:(length(r)+1)){
        #Caminata aleatoria
        S[j]=S[j-1]+r[j-1]

        #Proporciones Kelly Bayesiano
        fb[j]=(2*S[j]-j+1+a-b)/(j-1+a+b)
}

K=rep.int(1,length(r)+1)
Kb=rep.int(1,lenght(r)+1)
Km=rep.int(1,length(r)+1)

for(j in 2:(length(r)+1)){
```

```
            #Capital Kelly No Bayesiano
            K[j]=K[j-1]*(1+f)^(r[j-1])*(1-f)^(1-r[j-1])

            #Capital Kelly Bayesiano
            Kb[j]=Kb[j-1]*(1+fb[j-1])^(r[j-1])*(1-fb[j-1])^(1-r[j-1])

            #Capital Media A Priori
            Km[j]=Km[j-1]*(1+fp)^(r[j-1])*(1-fp)^(1-r[j-1])
      }
```

Bibliografía

[B] Breiman L. (1961), *Optimal Gambling Systems for Favorable Games*, California: Fourth Berkeley Symposium on Probability and Statistics.

[BW] Browne S., Whitt W. (1996), *Portfolio Choice and the Bayesian Kelly Criterion*, Advances in Applied Probability, 28 (4).

[C] Calva D. (2016), *Una introducción al criterio de Kelly y su Aplicación a Juegos Especulativos y de Mercado*, Tesis de licenciatura, Universidad Nacional Autónoma de México, 1 ,20-24.

[K] Kelly J. L. (1956), *A New Interpretation of Information Rate*, Nueva York: Bell System Technical Journal.

[L] Lauritzen S. (2007), *Exchangeability and de Finetti's Theorem*, Oxford: University of Oxford, Recuperado el 1 de enero de 2019, http://www.stats.ox.ac.uk/ steffen/teaching/grad/definetti.pdf.

[O] Ortega J. (2018), *Probabilidad Avanzada: Maestría en Probabilidad y Estadística*, Guanajuato: Centro de Invastigación en Matemáticas, Recuperado el 29 de septiembre de 2019, https://www.cimat.mx/ jortega/MaterialDidactico/PAv18/Cap2.pdf

[P] Poundstone W. (2005), *Fortune's Formula: The Untold Story of the Scientific Betting System That Beat the Casinos and Wall Street*, Nueva York: Hill & Wang.

[R1] Rincón L. (2014), *Introducción a la Probabilidad*, México: Departamento de Matemáticas.

[R2] Rincón L. (2012), *Introducción a la Teoría del Riesgo*, México: Departamento de Matemáticas.

[R3] Rincón L. (2012), *Introducción a los Procesos Estocásticos*, México: Departamento de Matemáticas.

[R4] Rincón L. (2007), *Curso Intermedio de Probabilidad*, México: Departamento de Matemáticas, 1, 38-39.

[S] Shannon C. E. (1948), *A Mathematical Theory of Communication*, Nueva York: Bell System Technical Journal.

[T] THORP E. O. (1997), *The Kelly Criterion in BlackJack Sports Betting and the Stock Market*, California: Finding the Edge: Mathematical Analysis of Casino Games.