



Propuesta Para Obtener Distribuciones A Priori Para Los Parámetros De La Distribución Beta

**Luis Gabriel Arroyo Bravo
Fabian Alejandro Lasso Balanta**

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2020

Propuesta Para Obtener Distribuciones A Priori Para Los Parámetros De La Distribución Beta

Luis Gabriel Arroyo Bravo
Fabian Alejandro Lasso Balanta

Trabajo de grado presentado como requisito parcial para optar al título de:
Estadístico

Director:
Ph.D. Jose Rafael Tovar Cuevas

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2020

Dedicatoria

A mis dos madres, Olga y Agripina:

Por brindarme todo su amor, por su entrega y sacrificio para que fuera un hombre de principios y de valores. Por alegrarse en mis alegrías y sentir como suyas mis tristezas, por ser mi cielo y mi suelo.

A mi padre, Luis Carlos:

Por ser un caballero, por los consejos, porque de ellos aprendí a sobrepasar los obstáculos que se presentan en la vida. Por su apoyo incondicional, y sobre todo, por brindarme el más puro amor.

A mi esposa y mi hijo, Katherine y Stevan:

por ser mi inspiración, por compartirme de su amor infinito y hacer de mi un mejor ser humano. Por darle color a mis días grises, colmándome de felicidad en cada instante.

A mis tíos y primos:

Por protegerme, amarme y siempre creer en mí. Por alegrarse en cada victoria que logré en la vida y acompañarme en cada fracaso.

Luis Gabriel Arroyo Bravo

A mi madre y a mi hermana, Flor y Lina:

Por siempre brindarme su apoyo incondicional, su gran capacidad de siempre salir adelante ante cualquier adversidad, y el amor que me demuestran cada día. Son mi mayor motivación.

Fabian Alejandro Lasso Balanta

Agradecimientos

Agradecemos a José Rafael Tovar Cuevas, profesor de la Escuela de Estadística de la Universidad del Valle, por su acompañamiento incondicional en el desarrollo de la presente investigación. Por su forma de enseñar, por los conocimientos que aportó para nuestra formación profesional, y que en parte sirvieron para nuestro crecimiento personal.

Agradecemos a Miguel Ángel Marmolejo Lasprilla, profesor del Departamento de Matemáticas de la Universidad del valle, por compartir sus conocimientos y su disponibilidad de tiempo para la asesoría, que fue clave en el desarrollo del presente trabajo.

Agradecemos a la Universidad del Valle, nuestra alma mater, y en especial a la Escuela de Estadística por brindarnos la oportunidad de aprender esta maravillosa profesión.

Resumen

En la práctica, es común modelar a través de la distribución Beta, las variables cuyo dominio se encuentra dentro del intervalo $(0, 1)$, caso de las proporciones o porcentajes. En este trabajo, se propone una metodología para obtener distribuciones a priori para los parámetros de forma de la distribución Beta, lo cual conlleva un nuevo enfoque para dicha distribución desde el paradigma Bayesiano, pues, la mayoría de las veces es utilizada como distribución previa cuando los datos de la muestra pueden ser ajustados a una distribución binomial, pero en el caso de interés para este trabajo, los datos de la muestra pueden ser ajustados con una distribución Beta lo que implicaría tener una función de verosimilitud que es el producto de n densidades Beta. Se obtuvo inicialmente una distribución conjunta para los parámetros de la distribución a partir de la información elicitada para la media y la varianza de la distribución. Finalmente, las distribuciones a priori para los parámetros de forma fueron obtenidas marginalizando la distribución conjunta.

Palabras clave: Distribución Beta, distribución a priori, distribución conjunta, distribución marginal, Inferencia Bayesiana

Abstract

In practice, it is common to model through the Beta distribution, the variables whose domain is within the interval $(0, 1)$, in the case of proportions or percentages. In this article, a methodology is proposed to obtain a priori distributions for the shape parameters of the Beta distribution, which entails a new approach for these distribution from the Bayesian paradigm, since, most of the time, it is used as a prior distribution when the sample data can be adjusted to a binomial distribution, but in the case of interest for this work, the sample data can be adjusted with a Beta distribution which would imply having a likelihood function that is the product of n densities Beta. A joint distribution for the distribution parameters was initially obtained from the information elicited for the mean and variance of the distribution. Finally, the a priori distributions for the shape parameters were obtained marginalizing the joint distribution.

Keywords: Beta distribution, a priori distribution, joint distribution, marginal distribution, Bayesian inference

Contenido

Resumen	ix
Lista de Figuras	1
1 Introducción	3
2 Planteamiento de la investigación	5
2.1 Problema Estadístico	5
2.1.1 Justificación	6
2.1.2 Objetivos	7
3 Antecedentes	8
4 Marco Teórico	12
4.1 Distribución Beta	12
4.2 Regla de L'Hôpital	13
4.3 Regla de los trapecios	13
4.4 Distribución Kumaraswamy	14
4.5 Inferencia Bayesiana	16
4.5.1 Teorema de Bayes	16
4.5.2 Elicitación	16
4.6 Cadenas de Markov y Método Montecarlo (MCMC)	18
4.6.1 Muestreo de Gibbs	19
4.6.2 Algoritmo Metropolis-Hasting	19
5 Metodología	22
5.1 Tratamiento Bayesiano para Verosimilitudes Beta	22
6 Resultados	38
6.1 Distribución A Priori para α	38
6.2 Distribución A Priori para β	42
7 Conclusiones y Recomendaciones	43
Bibliografía	45

Lista de Figuras

5-1	Comportamiento de σ^2 vs Valores fijos de (α, β)	25
5-2	Comportamiento empírico de σ^2	25
5-3	Gráfica de σ^2 dado los valores de μ	27
6-1	Curva de densidad de la distribución a priori de μ	38
6-2	Curva de densidad de la distribución a priori aproximada con el algoritmo .	39
6-3	Densidad de números aleatorios obtenidos para distribución a priori de α . .	41
6-4	Densidad de números aleatorios obtenidos para distribución a priori de β . .	42

Declaración

Nos permitimos afirmar que hemos realizado el presente Trabajo de Grado de manera autónoma y con la única ayuda de los medios permitidos y no diferentes a los mencionados en el propio trabajo. Todos los pasajes que se han tomado de manera textual o figurativa de textos publicados y no publicados, los hemos reconocido. Ninguna parte del presente trabajo se ha empleado en ningún otro tipo de Tesis o Trabajo de Grado.

Igualmente declaramos que los datos utilizados en este trabajo están protegidos por las correspondientes cláusulas de confidencialidad.

Santiago de Cali, 12.01.2020

(Luis Gabriel Arroyo Bravo)

(Fabian Alejandro Lasso Balanta)

1 Introducción

La distribución Beta, en la práctica ha cobrado mucha importancia debido a que su comportamiento permite modelar variables cuyo rango de valores se encuentra restringido en un intervalo de números reales con longitud finita. En especial, cuando se desea modelar variables cuyos valores se encuentran definidos en el espacio $(0, 1)$, como lo son las tasas o las proporciones, entre otros.

En la actualidad, las investigaciones en las cuales se trabaja con esta distribución son escasas, y al momento de realizar las estimaciones pertinentes para obtener resultados, los investigadores generalmente se apoyan en el paradigma clásico de la estadística. Por lo tanto, solo recurren a los ya conocidos métodos de estimación como el de los momentos y el de máxima verosimilitud. Sin embargo, cuando se desea obtener información sobre la distribución Beta planteada como verosimilitud, desde el enfoque Bayesiano de la estadística, es realmente muy poco lo que se encuentra debido a la complejidad y limitaciones que este modelo presenta.

El presente trabajo, plantea una propuesta metodológica para obtener distribuciones a priori para los parámetros de forma de la distribución beta, lo cual conlleva un nuevo enfoque para dicha distribución desde el paradigma Bayesiano. Así, se abordan las principales limitaciones que se tienen actualmente para el mencionado objetivo, desde la poca información literaria, hasta el hecho de que estos parámetros no tienen interpretación práctica.

Teniendo en cuenta lo anterior, y además, que el pilar fundamental de la estadística Bayesiana es la inclusión de creencias subjetivas sobre los parámetros de interés mediante el proceso de elicitación, en primera instancia se plantea realizar dichos procesos entorno a la media y la varianza de la distribución beta. Es decir, que en principio estas medidas serán los parámetros de interés.

Por ello, se asignan distribuciones a priori que consideran los espacios paramétricos y las restricciones matemáticas y/o estadísticas que pueden llegar a existir entre la media y la varianza. Esto, con la finalidad de obtener distribuciones a priori para los parámetros de forma, mediante las técnicas que la estadística ofrece para ello, como las transformaciones bivariadas de variables aleatorias.

En consecuencia, se llega a una expresión matemática para la distribución a priori conjunta de los parámetros de forma, y marginalizando se consiguen las a priori marginales para los mismos. Específicamente, una de esas distribuciones a priori se considera no informativa y es de forma cerrada. Además, se emplea el método de la transformada inversa para obtener números aleatorios que provengan de esta última distribución.

Del mismo modo, se consigue la forma de una distribución a priori marginal informativa para el parámetro restante. No obstante, ésta no tiene la característica de ser de forma cerrada, lo que conlleva a múltiples dificultades metodológicas desde el contexto bayesiano. Sin embargo, se proponen soluciones algorítmicas para aproximar dicha densidad, y que en conjunto con los métodos de Cadenas de Markov y Montecarlo, sean de utilidad para obtener números aleatorios de dicha distribución.

Finalmente, se realizan una serie de conclusiones y recomendaciones, para que en trabajos posteriores que den continuidad a esta metodología, se pueda derivar en el proceso de estimación y evaluación de la misma.

2 Planteamiento de la investigación

2.1. Problema Estadístico

En la práctica, muchas de las variables aleatorias que se estudian, se mueven en espacios acotados, de allí que encontrar distribuciones probabilísticas que modelen de forma adecuada dichos fenómenos cobra mucha importancia. Una alternativa para lo anterior es la distribución beta, pues según Canavos (1998): “ Se ha utilizado para representar variables físicas cuyos valores se encuentran restringidos a un intervalo de longitud finita y para encontrar ciertas cantidades que se encuentran como límites de tolerancia sin necesidad de la hipótesis de una distribución normal”(p.147).

Específicamente, cuando se trata de modelar proporciones, indicadores de desarrollo y otros fenómenos cuyos rangos de movimiento están definidos en el espacio $(0, 1)$, se emplea el modelo probabilístico Beta estándar, el cual es un caso especial del modelo Beta generalizado, y es el que se tratará en el presente trabajo de grado. Actualmente, la mayoría de los autores que trabajan con esta distribución, se apoyan en la investigación y resultados que se han obtenido en ella desde el paradigma de la estadística clásica. Por ende, al realizar estimaciones puntuales de los parámetros, recurren a los ya conocidos métodos de los momentos y de máxima verosimilitud. En este último, se apoyan en métodos iterativos como el algoritmo Newton Rapson, pues a la hora de derivar la función log-verosimilitud con respecto a los parámetros, la expresión se indetermina. En consecuencia, si se presentan problemas tales como tamaños muestrales muy pequeños o que la muestra que se tenga a disposición no este aleatorizada, se tendrán inconvenientes a la hora de sacar conclusiones en la investigación, pues estas carecerán de validez externa.

Este tipo de validez, está relacionada con la selección de las unidades estadísticas que serán medidas o encuestadas, aquí la muestra debe ser representativa de la población, de la selección dada y para ello se necesita que la forma como ella sea seleccionada no tenga sesgos, lo cual se garantiza mediante el uso del muestreo probabilístico. (Klinger, 2011, p.26).

En efecto, si se tienen dificultades en el paradigma clásico o simplemente se desea contar con una fuente de estudio adicional para concluir, el paradigma bayesiano surge como una alternativa importante. Tener estimaciones bayesianas para cualquier caso de investigación es beneficioso, pues esta enriquece el análisis, dado que los estimadores que se obtienen

consideran el conocimiento o experiencia subjetiva que tienen los expertos en el problema objeto de estudio. Sin embargo, si se quiere profundizar en la inferencia Bayesiana para esta distribución, se encuentran muchas limitaciones, ya que hasta el momento en la literatura no se cuenta con una distribución a priori para los parámetros de la Beta estándar, entre otras causas porque no se les ha dado interpretación alguna.

2.1.1. Justificación

En la literatura se encuentran trabajos en los que se utiliza la distribución Beta, ya que dada su naturaleza, muchas de las variables que se estudian se mueven en el espacio $(0, 1)$. Pero en la mayoría de éstos no se estiman los parámetros de forma bayesiana, debido a que la forma analítica de la distribución es muy compleja y se hace difícil elicitar/ hallar una distribución a priori para asociarla a dicha verosimilitud y obtener resultados.

Esto puede derivar en un problema, pues se pueden distorsionar las conclusiones del estudio, dado que al basarse solamente en el enfoque clásico, existe la posibilidad de que se presenten dificultades en los métodos de estimación, ya sea por una muestra pequeña y/o no aleatoria o por los algoritmos de iteración con los que cuenta el mencionado enfoque a la hora de estimar los parámetros.

Una razón pragmática para el uso de los métodos bayesianos es la flexibilidad inherente introducida por su incorporación de múltiples niveles de aleatoriedad y la capacidad resultante de combinar información de diferentes fuentes, al tiempo que incorpora todas las fuentes razonables de incertidumbre en los resúmenes inferenciales. Tales métodos conducen naturalmente a estimaciones suavizadas en estructuras de datos complicadas y, por consiguiente, tienen la capacidad de obtener mejores respuestas del mundo real. (Gelman et al. , 2014 , p.24)

Por ende, ya que a los parámetros de la distribución Beta no se les ha encontrado una interpretación, se hace necesario encontrar una distribución a priori para los momentos, pues estos tienen interpretación práctica, además de estar en función de los parámetros de forma en cuestión. Así, por medio de conceptos probabilísticos obtener una distribución a priori para estos últimos, que permita asociar la verosimilitud con el conocimiento que se tenga acerca del problema de estudio para conseguir una distribución posterior, con la cual se puedan realizar estimaciones bayesianas y finalmente superar dicha dificultad.

2.1.2. Objetivos

Objetivo General

- Proponer una distribución a priori para los parámetros de la Distribución Beta.

Objetivos Específicos

- Proponer una distribución a priori para la media y la varianza de la Distribución Beta.
- Proponer una distribución a priori conjunta para la media y la varianza de la Distribución Beta.
- Proponer una distribución a priori para los parámetros de forma de la Distribucion Beta, a través de la distribución conjunta planteada para la media y varianza.

3 Antecedentes

Ma y Leijon (2011), tuvieron como principal objetivo proponer un algoritmo basado en el marco de inferencia variacional para aproximar la distribución posterior de los parámetros de la distribución Beta. Plantean que realizar la estimación desde el enfoque bayesiano es realmente complejo, especialmente en los Modelos de Mixtura Beta(BMM), pues la expresión de la distribución posterior que resulta al tratar de estimar los parámetros, utilizando la conjugada de esta distribución como a priori, es analíticamente intratable. Por tanto, bajo la inferencia variacional proponen un método más flexible, el cual consiste en una aproximación a la distribución a priori de los parámetros, y de esta manera obtener una solución de forma cerrada.

Así, emplearon principios del método de aproximación factorizada, para asumir independencia entre los dos parámetros de forma. En este sentido, asignan una distribución a priori Gamma para cada parámetro, y al multiplicarlas obtienen la distribución a priori conjunta de los mismos. Ma y Leijon decidieron emplear un modelo jerárquico para realizar la estimación, estableciendo una variable indicadora para cada observación de la verosimilitud, y asumiendo que los parámetros de dicha variable, se distribuyen la Dirichlet bajo el paradigma bayesiano.

Después de esto, plantean una distribución conjunta que involucra la verosimilitud y las variables indicadoras, condicionada a todo el conjunto de parámetros e hiperparámetros involucrados en el contexto del problema. Lo anterior, con la finalidad de obtener el logaritmo de esa expresión y así seguir el esquema de actualización de la inferencia variacional para estimar los hiperparámetros. Proponen un método iterativo para calcular el valor esperado del logaritmo con respecto a cada parámetro, condicionado a todo el grupo restante de variables, parámetros e hiperparámetros.

Considerando un parámetro de la Beta como variable de interés, y partiendo del valor esperado calculado, los autores llegan a una solución óptima y tratable para la distribución a priori y posteriori del mismo, con sus respectivos hiperparámetros. Análogamente, obtienen una solución óptima para el parámetro restante de la distribución Beta.

Por último, Ma y Leijon verificaron con datos simulados la eficiencia y la corrección del método propuesto. Concluyen que en comparación con el algoritmo de maximización

del valor esperado (EM) convencional, el método bayesiano que propusieron, evita los inconvenientes del sobreajuste y del cálculo numérico iterativo en la etapa de maximización. También, sugieren que este método podría estimar una distribución más precisa, cuando la cantidad de datos es pequeña. A medida que aumenta el número de muestras, el método Bayesiano converge al mismo resultado que el algoritmo EM. Para la evaluación de datos reales, aplicaron el BMM Bayesiano propuesto para la tarea de modelado de imágenes SAR, detección de color de piel y cuantificación de parámetros de LSF. Determinaron que el BMM bayesiano propuesto era superior al de modelos de mixtura Gaussiana, debido al soporte de los límites de los datos y la flexibilidad del marco bayesiano.

Kakamu y Nishino (2018), realizaron un estudio para estimar los parámetros de la distribución Beta Generalizada (BG) basado en datos agrupados desde una perspectiva Bayesiana utilizando el algoritmo bloques aleatorios a medida Metrópolis-Hastings (TaRBMH), que fue propuesto por Chib y Ramamurthy (2010) apoyados en el ya existente algoritmo de caminata aleatoria Metrópolis-Hastings (RWMH).

En este nuevo algoritmo, los parámetros son codificados y se agrupan en varios bloques distintos aleatoriamente, lo cual evita una mala elección de bloques a priori. Luego, cada bloque de parámetros se actualiza en secuencia por un paso del Metrópolis-Hastings teniendo en cuenta la densidad propuesta, hasta que ésta se adapta tanto en ubicación como en la curvatura de la densidad posterior del bloque, condicionado al valor más actual de los parámetros en el Bloques restantes.

Inicialmente, los autores tienen en cuenta la verosimilitud Beta Generalizada con sus respectivos parámetros, a los cuales les asigno una distribución a priori a cada uno de ellos teniendo en cuenta las restricciones que éstos presentan. Al parámetro de localización se le asignó una distribución a priori Normal, al parámetro de escala y a los dos de forma les asignaron distribuciones a priori Gamma, y cada una de estas distribuciones cuentan con sus correspondientes hiperparámetros. Luego, multiplicando estas distribuciones a priori con la verosimilitud BG obtienen la función de densidad de la distribución posterior e implementaron el algoritmo TaRBMH para producir las estimaciones posteriores.

Finalmente, Kakamu y Nishino realizaron una prueba con datos simulados y confirmaron que el algoritmo que TaRBMH realiza mejores estimaciones que el algoritmo RWMH. También, lo probaron con una base de datos reales de los ingresos familiares en Estados Unidos cuya variable de interés era la proporción salarial respecto al ingreso base, y desarrollando la estimación bayesiana por el algoritmo TaRBMH, comprobaron que estas siguen siendo más precisas que las realizadas por el método RWMH. Además, enfatizaron que en iteraciones más cortas del algoritmo RWMH se podría conducir a una estimación inexacta de los parámetros verdaderos, mientras que el algoritmo TaRBMH puede evitar

este problema, incluso para iteraciones más cortas.

Owen (2008), realizó la estimación de los parámetros de la distribución Beta, mediante los ya conocidos métodos de Máxima Verosimilitud y el de los Momentos, con el fin de examinar ambos métodos y luego compararlos con otros tres métodos propuestos de estimación de parámetros. Estos son:

1) El método Técnica de Revisión y Evaluación del Programa (PERT), donde se basa en (Farnum and Stanton 1987; Johnson 1997) para calcular un coeficiente llamado el valor “más probable”, para luego, dependiendo del rango en que se encuentre dicho valor, se estiman la media y la varianza de la distribución Beta, las cuales se encuentran en función del mencionado valor. Finalmente, en un procedimiento análogo al método de los Momentos, iguala dichas estimaciones a la expresiones poblacionales y despeja el valor estimado de los parámetros.

2) Una modificación de la Distribución de Potencia de Dos Colas (TSP), donde plantea que el comportamiento de esta distribución, la cual es una modificación de la distribución Triangular, se puede aproximar al de la distribución Beta, con la diferencia de que sus parámetros tienen una mejor interpretación en la práctica. Por ende, utilizando el método de Máxima Verosimilitud, estima los parámetros de esta y así obtiene una estimación para la varianza y la media de la distribución de Potencia de Dos Colas. Por último, dada las semejanza de las distribuciones, iguala estas estimaciones a las expresión poblacional de la varianza y la media de la distribución Beta, obteniendo así las estimaciones paramétricas para esta última.

3) Un estimador de cuantiles basado en el primer y tercer cuartil de la distribución Beta, en el cual propone un método empírico apoyado en el software estadístico R, que es establecer el primer y tercer cuartil poblacional de la distribución Beta, pues conoce el valor real de los parámetros. Acto seguido, mediante escenarios de simulación genera distintas muestras de esta distribución, donde el conjunto de valores paramétricos corresponden a un intervalo alrededor de la estimación de los parámetros obtenida por el método de los Momentos (recomienda que dicho intervalo no se aleje en más de una unidad de la estimación, para los dos límites), para después calcular de forma muestral los cuantiles referidos anteriormente. La estimación por este método, corresponde al conjunto de estimaciones que más se acercó al valor real de los cuantiles en cuestión. Finalmente, para un conjunto de datos reales, donde los parámetros son desconocidos, recomienda tomar como valor real de estos en la simulación, la estimación alcanzada por el método de los Momentos.

Owen aplicó estas técnicas de estimación a dos conjuntos de datos para ver qué tan bien se aproximan a los datos reales de las Grandes Ligas de Béisbol (proporción de bateo) y el

Departamento de Energía de los Estados Unidos (exposición a la radiación). Al final, obtuvo que el estimador de Cuantiles, el de Máxima Verosimilitud y el método de los Momentos funciona bien sin importar la cantidad de datos que se tengan en una muestra indexada por la distribución Beta, en contraste con los otros métodos. Sin embargo, Los estimadores PERT y TSP se adaptan mejor para muestras pequeñas de distribuciones Beta, aunque sus resultados no superan a los obtenidos mediante los tres métodos citados con anterioridad.

4 Marco Teórico

4.1. Distribución Beta

La familia de distribuciones beta está compuesta por todas las distribuciones con función de densidad de la forma:

$$f_Y(y|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \frac{(y-a)^{\alpha-1}(b-y)^{\beta-1}}{(b-a)^{\alpha+\beta-1}}, \quad a \leq y \leq b \quad (4-1)$$

Con $\alpha > 0$, $\beta > 0$. Esta se denota como $\text{Beta}(\alpha, \beta)$. Esta es reconocida como la distribución de Pearson Tipo I o II. Si $\beta = 1$, la distribución algunas veces es llamada Power-function. (Johnson et al. 1995)

La forma acumulada viene dada por

$$F_Y(y|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_a^y \frac{(t-a)^{\alpha-1}(b-t)^{\beta-1}}{(b-a)^{\alpha+\beta-1}} dt \quad (4-2)$$

Si se realiza la transformación

$$X = \frac{Y-a}{b-a}$$

Se obtiene la función de densidad

$$f_x(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 \leq x \leq 1 \quad (4-3)$$

Esta es la forma estándar de la distribución Beta con parámetros α, β . (Johnson et al. 1995)

Se sabe que la forma de la distribución Beta depende de los valores que tomen los parámetros α, β . Además, que

$$B_x(\alpha, \beta) = \int_0^x t^{\alpha-1} (1-t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad (4-4)$$

Donde $\Gamma(v)$ hace referencia a la función Gamma $\Gamma(v) = \int_0^\infty t^{v-1}e^{-v}dv$, para algún v .

Si X se distribuye Beta estándar, el r -ésimo momento entorno a cero es

$$\mu'_r = \frac{B(\alpha + r, \beta)}{B(\alpha, \beta)} = \frac{\Gamma(\alpha + r)\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\alpha + \beta + r)} = \frac{\alpha^{[r]}}{(\alpha + \beta)^{[r]}} \quad (4-5)$$

(si r es un entero), donde $y^{[r]} = y(y + 1)\dots(y + r - 1)$ es ascendentemente factorial. En particular (Johnson et al. 1995).

$$\mu = \frac{\alpha}{\alpha + \beta} \quad ; \quad \sigma^2 = \alpha\beta(\alpha + \beta)^{-2}(\alpha + \beta + 1)^{-1} \quad (4-6)$$

4.2. Regla de L'Hôpital

En términos de límites de las funciones matemáticas, existen formas del tipo $0/0$ o ∞/∞ , las cuales son llamadas indeterminadas. Nombradas de esta manera, debido a que no garantizan que el límite existe.

Para encontrar el límite de estas formas indeterminadas, se puede usar el teorema llamado la regla de L'Hôpital. Este teorema establece que bajo ciertas condiciones el límite del cociente $f(x)/g(x)$ es determinado por el límite del cociente de las derivadas. (Larson et al. 2006)

$$\frac{f'(x)}{g'(x)} \quad (4-7)$$

El teorema establece lo siguiente: Sean f y g funciones que son derivables en el intervalo abierto (a, b) que contiene un valor c . Se debe asumir que $g'(x) \neq 0$ para todo x en (a, b) , excepto posiblemente el propio c . Si el límite de $f(x)/g(x)$ cuando x tiende a c produce la forma indeterminada $0/0$, entonces

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} \quad (4-8)$$

Este resultado también aplica si el límite de $f(x)/g(x)$ cuando x tiende a c produce cualquiera de las formas indeterminadas ∞/∞ , $-\infty/\infty$, $\infty/-\infty$, o $-\infty/-\infty$. (Larson et al. 2006)

4.3. Regla de los trapecios

En conceptos matemáticos, algunas funciones elementales simplemente no tienen antiderivadas o primitivas que sean funciones elementales. Si se ha de calcular una

integral definida cuyo integrando no admite antiderivada, se debe recurrir a un método de aproximación como lo es la regla de los trapecios, el cual es un método de integración numérica que permite aproximar el valor de la integral definida. (Larson et al. 2006)

El teorema propone, que si una función f es continua y positiva en el intervalo $[a, b]$ se debe dividir el intervalo en n subintervalos con un ancho $\Delta x = (b - a)/n$, de tal modo que $a = x_0 < x_1 < \dots < x_n = b$. Acto seguido, se forma un trapecio para cada subintervalo, donde el área del i -ésimo trapecio es construido de la siguiente manera.

$$\left[\frac{f(x_{i-1}) + f(x_i)}{2} \right] \left(\frac{b - a}{n} \right) \quad (4-9)$$

Para luego, sumar todas las áreas de los n trapecios y finalmente tomar el límite cuando $n \rightarrow \infty$ con la intención de obtener el valor aproximado de la integral $\int_a^b f(x)dx$, donde la suma del área de todas las áreas de los n trapecios se muestra a continuación.

$$\frac{b - a}{2n} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + 2f(x_n)] \quad (4-10)$$

Y el límite de esta función cuando $n \rightarrow \infty$ es

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{b - a}{2n} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + 2f(x_n)] \\ &= \lim_{n \rightarrow \infty} \frac{[f(a) - f(b)](b - a)}{2n} + \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x \\ &= 0 + \int_a^b f(x)dx \end{aligned} \quad (4-11)$$

Entonces, la regla de los trapecios para aproximar $\int_a^b f(x)dx$ está dada por

$$\int_a^b f(x)dx \approx \frac{b - a}{2n} [f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + 2f(x_n)] \quad (4-12)$$

Además, como $n \rightarrow \infty$, el lado derecho se aproxima a $\int_a^b f(x)dx$. (Larson et al. 2006)

4.4. Distribución Kumaraswamy

La distribución Kumaraswamy pertenece a la familia de distribuciones de probabilidad continua definidas en el intervalo $(0,1)$. Esta distribución fue propuesta originalmente por Poondi Kumaraswamy en 1980, cuyos parámetros de forma (a_1, b_1) corresponden al conjunto

de los reales no negativos, y su forma es similar a la distribución Beta. Sin embargo, es más simple de implementar, especialmente en estudios de simulación, ya que su función de densidad de probabilidad, función de distribución acumulativa y funciones cuantiles se pueden expresar en forma cerrada. (Jones 2009)

A continuación, en la ecuación 4-13 se muestra la forma de su distribución de densidad, y en 4-14 su función de distribución acumulada. Posteriormente, se puede observar lo que se debe tener en cuenta al momento de obtener la media y la varianza de esta distribución. (Jones 2009)

$$f(x|a_1, b_1) = a_1 b_1 x^{a_1-1} (1 - x^{a_1})^{b_1-1} \quad ; 0 < x < 1; \quad a_1 > 0, \quad b_1 > 0 \quad (4-13)$$

$$F(x) = \int_0^x f(t)dt = 1 - (1 - x^{a_1})^{b_1} \quad (4-14)$$

Ahora bien, para obtener la media y la varianza, se debe tener en cuenta los momentos de esta distribución, donde el momento n viene denotado por:

$$m_n = \frac{b_1 \Gamma(1 + \frac{n}{a_1}) \Gamma(b_1)}{\Gamma(1 + b_1 + \frac{n}{a_1})} \quad (4-15)$$

De este modo, μ^* y σ^{2*} son representadas por las ecuaciones 4-16 y 4-17 respectivamente. Además, su mediana viene dada por la ecuación 4-18.

$$\mu^* = m_1 = \frac{b_1 \Gamma(1 + \frac{1}{a_1}) \Gamma(b_1)}{\Gamma(1 + b_1 + \frac{1}{a_1})} \quad (4-16)$$

$$\sigma^{2*} = m_2 - m_1^2 \quad (4-17)$$

$$m_e^* = (1 - 2^{-1/b_1})^{1/a_1} \quad (4-18)$$

Donde $\Gamma(v)$ hace referencia a la función Gamma $\Gamma(v) = \int_0^\infty t^{v-1} e^{-t} dt$, para algún v . (Jones 2009).

4.5. Inferencia Bayesiana

La inferencia Bayesiana hace uso de toda la información de los datos, además, da solución a algunos problemas incómodos sobre la teoría del muestreo en cuanto a la elección de estimadores e intervalos de confianza, y proporciona una manera satisfactoria de introducir y mantener explícitamente los supuestos sobre el conocimiento previo o la ignorancia. Se debe reconocer que algún conocimiento previo se emplea en todos los sistemas inferenciales. (Box y Tiao 1992)

4.5.1. Teorema de Bayes

Suponga que $y = (y_1, \dots, y_n)$ es un vector de n observaciones cuya distribución de probabilidad $f(y|\theta)$ depende de los valores de k parámetros $\theta = (\theta_1, \dots, \theta_k)$. Suponga también que θ tiene una distribución de probabilidad $\pi(\theta)$. Entonces, dado los datos observados y , la distribución condicional de θ es

$$f(\theta|y) = \frac{L(\theta|y)\pi(\theta)}{\int_{\Theta} L(\theta|y)\pi(\theta)d\theta} \quad (4-19)$$

Donde Θ es el espacio paramétrico de θ , y

$$Ef(y|\theta) = c^{-1} = \int_{\Theta} L(\theta|y)\pi(\theta)d\theta \quad (4-20)$$

Donde la integral está definida sobre todo el espacio paramétrico de θ , y donde $Ef(y|\theta)$ es el valor esperado de $f(\theta)$ con respecto a la distribución $\pi(\theta)$. La cantidad c es meramente una constante de normalización necesaria para asegurar que la distribución a posteriori $f(\theta|y)$ se integre o sume a uno. Así, se puede escribir 4-19 alternativamente como (Box y Tiao 1992).

$$f(\theta|y) = cL(\theta|y)\pi(\theta). \quad (4-21)$$

La declaración de (4-19) o su equivalente (4-21) es usualmente referido como el Teorema de Bayes. En esta expresión, $\pi(\theta)$, dice lo que se sabe sobre θ sin el conocimiento de los datos, es llamada distribución a priori de θ . Correspondientemente, $f(\theta|y)$, expresa lo que se sabe sobre θ dado el conocimiento de los datos, es llamada distribución a posteriori de θ . (Box y Tiao 1992).

4.5.2. Elicitación

Para Garthwaite, P., Kadane, J. y O'hagan, A. (2005), la elicitación es el proceso de formular el conocimiento y las creencias de una persona acerca de una o más cantidades inciertas (parámetros) en una distribución de probabilidad (conjunta) para esas cantidades.

En el contexto del análisis estadístico bayesiano, la elicitación de datos con mayor frecuencia surge como un método para especificar la distribución a priori de uno o más parámetros desconocidos de un modelo estadístico. En este contexto, la distribución a priori se combinará con la probabilidad a través del teorema de Bayes para derivar la distribución a posteriori. Sin embargo, este no es el único contexto en el que la elicitación es importante.

Es conveniente pensar que la tarea de elicitación involucra a un facilitador, que ayuda al experto a formular los conocimientos en forma probabilística. En el contexto de obtener una distribución previa para un análisis Bayesiano, es el conocimiento previo del experto el que se obtiene, pero en general, el objetivo es expresar el conocimiento actual del experto en forma probabilística. Si el experto es un estadístico o está muy familiarizado con los conceptos estadísticos, entonces puede que no haya una necesidad formal de un facilitador, pero esto es raro en la práctica. La elicitación es un proceso complejo que exige una variedad de habilidades si se va a hacer bien, y el rol de facilitador es importante. (Garthwaite et al. 2005)

Los métodos de obtención más complejos generalmente imponen una estructura a la opinión de un experto al suponer que su conocimiento puede estar bien representado por algún miembro de una familia específica de distribuciones. Si el experto ha especificado información que se ajusta a una distribución paramétrica conveniente, entonces tiene sentido usarla. Esta estrategia tiene costos y beneficios similares a los de los modelos estadísticos en general. Los miembros de la familia hipotética se distinguen por parámetros (llamados hiperparámetros), y la tarea de obtención se reduce a la elección de valores de hiperparámetro apropiados para capturar las características principales de la opinión del experto. Si la opinión del experto no corresponde aproximadamente a ningún miembro de la familia, es probable que aparezcan discrepancias en las respuestas del experto a las preguntas sobre elicitación, así como un modelo de muestreo puede sucumbir a los controles de diagnóstico tradicionales. La familia de distribuciones se elige típicamente para ser la familia de conjugados naturales (o una extensión manejable de esa familia), que facilita el análisis posterior si se dispone de datos de muestra, aunque los avances en el cálculo bayesiano a través del método Markov Chain Monte Carlo (MCMC) hace viable el uso de otras familias. Para muchos modelos de muestreo, la familia de conjugados es razonablemente flexible y puede representar una variedad de opiniones a través de la elección adecuada de hiperparámetros. Mayor flexibilidad está disponible a través de mezclas de conjugados. (Garthwaite et al. 2005)

Un sistema de declaraciones de probabilidad, las cuales se pretende elicitar, es coherente si todas las probabilidades son consistentes con las leyes de probabilidad. Por ejemplo, un experto afirma que $P(E) = 0,4$, $P(F) = 0,3$, y $P(EoF) = 0,6$, cuando E y F son eventos mutuamente excluyentes, entonces estas probabilidades no son coherentes. Una forma de verificar la calidad de las declaraciones de un experto es que el facilitador solicite conjuntos de evaluaciones de probabilidad que permitan pruebas de coherencia. Debemos

esperar que las declaraciones obtenidas por el experto no pasen las pruebas de coherencia. Esto es casi inevitable en vista de la imprecisión con la que el experto puede hacer estos juicios. (Garthwaite et al. 2005)

Sin embargo, tenga en cuenta que las declaraciones inconsistentes del experto pueden simplemente indicar que la distribución del experto no puede ser representada adecuadamente por un miembro de la familia elegida. Por ejemplo, si se buscan las creencias del experto con respecto a una proporción π , entonces el facilitador podría elegir trabajar con el supuesto de una distribución Beta. Luego, si el experto especifica que la mediana es 0,4, el cuartil inferior es 0,3, y el cuartil superior es 0,5, ninguna distribución beta se ajustará a estas especificaciones. La distribución $Beta(4,733; 7,1)$ se ajusta a la mediana y al cuartil inferior pero tiene un cuartil superior de 0.494, mientras que la distribución $Beta(4,16; 6,24)$ se ajusta a la mediana y al cuartil superior, pero tiene un cuartil inferior de 0,293. Pero no es razonable esperar que el experto especifique estos cuantiles con tanta precisión, y las evaluaciones en la práctica no se verían como un desafío al supuesto de una distribución beta. (Garthwaite et al. 2005)

4.6. Cadenas de Markov y Método Montecarlo (MCMC)

Se han ideado muchos métodos inteligentes para construir y muestrear distribuciones posteriores arbitrariamente. La simulación de una cadena de Markov (también llamada cadena de Markov de Monte Carlo, o MCMC) es un método general basado en obtener valores de θ a partir de distribuciones aproximadas y luego corrigiendo dichos valores para aproximar mejor la distribución posterior objetivo, $p(\theta|y)$. (Gelman et al. 2014)

Se realiza un muestreo de forma secuencial, con la distribución de los valores muestreados en función del último valor obtenido; Por lo tanto, dichos valores forman una cadena de Markov. (Como se define en la teoría de la probabilidad, una cadena de Markov es una secuencia de variables aleatorias $\theta_1, \theta_2, \dots$, para la cual, para cualquier t , la distribución de θ^t dados todos los θ' s previos depende solo del valor más reciente, θ^{t-1} .) La clave del éxito del método, sin embargo, no es la propiedad de Markov sino que las distribuciones aproximadas se mejoran en cada paso de la simulación, en el sentido de que convergen a la distribución objetivo. (Gelman et al. 2014)

Considere una cadena de markov ergódica $(\theta^{(n)})_n \geq 0$ con estado de espacio $S \subset R^d$, kernel de transición $p(x, y)$ y distribución inicial $\pi^{(o)}$. La generación de un valor de esta cadena comienza con un valor para $\theta^{(o)}$ muestreado a partir de $\pi^{(o)}$. el valor de $\theta^{(1)}$ se distribuye con densidad $p(\theta^{(o)}, \cdot)$ y puede ser generados utilizando el método de la transformada inversa. Para $\theta^{(2)}$, este procedimiento es repetido por la distribución obtenida a partir de la densidad $p(\theta^{(1)}, \cdot)$, iterando este esquema mediante los pasos de la cadena se obtiene $\theta^{(n)}$ a partir de la

distribución con densidad $p(\theta^{(n-1)}, \cdot)$, para todo n . Como n tiende a ser grande, los valores se acercan cada vez más a los de la distribución límite π y pueden considerarse como valores aproximados de π . (Gamerman y Lopes 2006)

4.6.1. Muestreo de Gibbs

El muestreo de Gibbs es un esquema MCMC donde el kernel de transición es formado con la distribución condicional completa. Asuma que la distribución de interés es $\pi(\theta)$ donde $\theta = (\theta_1, \dots, \theta_d)'$. Cada uno de los componentes θ_i puede ser un escalar, un vector o una matriz. Considere también que la distribución condicional completa $\pi_i(\theta_i) = \pi(\theta_i | \theta_{-i})$, $i = 1, \dots, d$ están disponibles. Esto significa que son completamente conocidos y pueden ser muestreados. (Gamerman y Lopes 2006)

El problema a resolver es obtener valores de π cuando los esquemas de generación directa son costosos, complicados o simplemente no están disponibles pero cuando la generación de los π_i es posible. El Muestreo de Gibbs provee un esquema de generación alternativo basado en generaciones sucesivas a partir de la distribución condicional completa. Este puede ser descrito de la siguiente manera: (Gamerman y Lopes 2006)

1. Inicializar el contador de iteraciones de la cadena $j = 1$ y un conjunto de valores iniciales $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})'$;

2. Obtenga el nuevo valor $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_d^{(j)})'$ a partir de $\theta^{(j-1)}$ mediante generación sucesiva de valores

$$\begin{aligned}\theta_1^{(j)} &\sim \pi(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)}), \\ \theta_2^{(j)} &\sim \pi(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_d^{(j-1)}), \\ &\vdots \\ \theta_d^{(j)} &\sim \pi(\theta_d | \theta_1^{(j)}, \dots, \theta_{d-1}^{(j)}); \end{aligned}$$

3. Cambiar el contador j a $j + 1$ y retornar al paso 2 hasta alcanzar la convergencia.

Cuando se alcanza la convergencia, el valor resultante $\theta^{(j)}$ es obtenido a partir de π . Como el número de iteraciones incrementa, la cadena se acerca a su condición de equilibrio. (Gamerman y Lopes 2006)

4.6.2. Algoritmo Metropolis-Hasting

Considere una distribución π de la cual debe extraerse una muestra vía Cadenas de Markov. Nuevamente, vale la pena señalar que esta tarea solo tendrá sentido si la generación no

iterativa de π es muy complicada o costosa. En este caso, un kernel de transición $p(\theta, \phi)$ debe ser construido de manera tal que π sea la distribución de equilibrio de la cadena. Una forma simple de hacer esto es considerar cadenas reversibles donde el kernel p satisface (Gamerman y Lopes 2006)

$$\pi(\theta)p(\theta, \phi) = \pi(\phi)p(\phi, \theta), \quad \forall(\theta, \phi). \quad (4-22)$$

La ecuación (4-22) es también llamada como la ecuación de balance detallado. Aunque esta no es una condición necesaria para la convergencia, es una condición suficiente para que π sea la distribución de equilibrio de la cadena.

El kernel $p(\theta, \phi)$ consiste de 2 elementos: un kernel de transición arbitrario $q(\theta, \phi)$ y una probabilidad $\alpha(\theta, \phi)$ tal que: (Gamerman y Lopes 2006)

$$p(\theta, \phi) = q(\theta, \phi)\alpha(\theta, \phi), \quad \text{si } \theta \neq \phi$$

Así que, el kernel de transición define una densidad $p(\theta, \cdot)$ para cada posible valor de los diferentes parámetros de θ . Consecuentemente, hay una probabilidad positiva de que la cadena permanezca en θ dada por

$$p(\theta, \theta) = 1 - \int q(\theta, \phi)\alpha(\theta, \phi)d\phi$$

Estas dos formas pueden ser agrupadas en la expresión general

$$p(\theta, A) = \int_A q(\theta, \phi)\alpha(\theta, \phi)d\phi + I(\theta \in A) \left[1 - \int q(\theta, \phi)\alpha(\theta, \phi)d\phi \right] \quad (4-23)$$

para algún subconjunto A del espacio paramétrico. Así que, el kernel de transición define una mezcla de distribuciones para el nuevo estado ϕ de la cadena. Para $\phi \neq \theta$, esta distribución tiene una densidad y para $\phi = \theta$, esta distribución tiene un átomo de probabilidad. (Gamerman y Lopes 2006)

Hastings(1970) propuso definir la probabilidad de aceptación de tal manera que cuando se combine con el kernel de transición arbitrario, este define una cadena reversible. La expresión más comúnmente citada para la probabilidad de aceptación es

$$\alpha(\theta, \phi) = \min \left\{ 1, \frac{\pi(\phi)q(\phi, \theta)}{\pi(\theta)q(\theta, \phi)} \right\}$$

Además, bajo una distribución propuesta independiente, ϕ es independiente al valor actual de la cadena de Markov en θ . Lo anterior, deriva en este caso, que la probabilidad de aceptación esté determinada por la siguiente expresión.(Bautista 2007)

$$\alpha(\theta, \phi) = \min \left\{ 1, \frac{\pi(\phi)h(\theta)}{\pi(\theta)h(\phi)} \right\} \quad (4-24)$$

Donde h es la distribución propuesta independiente. El algoritmo basado en cadenas con kernel de transición (4-23) y probabilidad de aceptación (4-24) será llamado como algoritmo Metropolis-Hastings. (Gamerman y Lopes 2006)

El kernel de transición propuesto q hasta ahora se ha mantenido arbitrario y, por lo tanto, es una herramienta flexible para la construcción del algoritmo. Roberts y Smith(1994) muestran que si q es irreducible y aperiódico y $\alpha(\theta, \phi) > 0$, para cada valor posible de (θ, ϕ) , entonces el algoritmo define una cadena irreducible y aperiódica con kernel de transición p dado por (4-23) y limitando la distribución de π (Gamerman y Lopes 2006)

En términos prácticos, simular un valor a partir de π usando una cadena de Markov definida por la transición (4-23) puede ser llevado a cabo de la siguiente manera:

1. Inicialice la iteración contando $j = 1$ y establezca un valor inicial arbitrario $\theta^{(0)}$.
2. Mover la cadena a un nuevo valor ϕ generado a partir de la densidad $q(\theta^{(j-1)}, \cdot)$.
3. Evalúe la probabilidad de aceptación del movimiento $\alpha(\theta^{(j-1)}, \phi)$. dado por (4-24) . Si el movimiento es aceptado, $\theta^{(j)} = \phi$. Si no es aceptado, $\theta^{(j)} = \theta^{(j-1)}$ y la cadena no se mueve.
4. Cambie el contador de (j) a $(j + 1)$ y retorne al paso 2 hasta alcanzar convergencia. (Gamerman y Lopes 2006)

El paso 3 es desarrollado después de la generación de una variable u uniforme cuantitativa independiente. Si $u \leq \alpha$, el movimiento es aceptado y si $u > \alpha$ el movimiento no es permitido. El kernel de transición q solamente define un posible movimiento que puede ser confirmado de acuerdo al valor de α . Por esta razón, q es generalmente llamado el kernel propuesto o densidad propuesta(condicional) cuando se ve como una densidad $q(\theta, \cdot)$ (condicional). (Gamerman y Lopes 2006)

5 Metodología

5.1. Tratamiento Bayesiano para Verosimilitudes Beta

Partiendo de la premisa de que se tiene un problema, el cual analíticamente está sujeto a la distribución beta, es decir, que todos los posibles valores de la variable aleatoria de interés, solo toman sentido en el espacio comprendido en el intervalo $(0, 1)$. Se tiene que la función de densidad de la misma, está indexada por la ecuación (5-1).

$$f_x(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 \leq x \leq 1; \quad \alpha > 0, \quad \beta > 0 \quad (5-1)$$

con

$$E(X) = \mu = \frac{\alpha}{\alpha + \beta} \quad ; \quad Var(X) = \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (5-2)$$

α y β son los parámetros de forma de dicha distribución. Como ya se ha mencionado, estos no gozan de una interpretación práctica, lo cual no permite un proceso de elicitación si se desea encontrar estimaciones para los mismos desde el paradigma Bayesiano. Esto supone un panorama complicado para obtener distribuciones a priori para los parámetros en cuestión.

Por el contrario, la media y la varianza de esta distribución si tienen una interpretación en el marco contextual de cualquier problema, pues son indicadores de tendencia central y dispersión de los datos respectivamente. Por ello, en la búsqueda de alternativas que permitan realizar estimaciones bayesianas para α y β , en primera instancia se alude a μ y σ^2 como los parámetros de interés, teniendo en cuenta que dependen de los parámetros de forma, para elicitar y asignarles distribuciones a priori, para después derivar en distribuciones a priori para α y β .

En este contexto, se hace necesario conocer una serie de requerimientos para crear una idea de cuales distribuciones serán razonables, para asignarlas como distribuciones a priori de μ y σ^2 . El primero de estos, es establecer en qué espacio de los números reales tienen sentido estos últimos.

Así, teniendo en cuenta que $\beta \in (0, \infty)$, se establece que:

$$\lim_{\beta \rightarrow \infty} \mu = \lim_{\beta \rightarrow \infty} \frac{\alpha}{\alpha + \beta} = 0 \quad (5-3)$$

$$\lim_{\beta \rightarrow 0} \mu = \lim_{\beta \rightarrow 0} \frac{\alpha}{\alpha + \beta} = 1 \quad (5-4)$$

También $\alpha \in (0, \infty)$, entonces:

$$\lim_{\alpha \rightarrow 0} \mu = \lim_{\alpha \rightarrow 0} \frac{\alpha}{\alpha + \beta} = 0 \quad (5-5)$$

Mientras tanto, por la regla de L'Hopital, se tiene que:

$$\lim_{\alpha \rightarrow \infty} \mu = \lim_{\alpha \rightarrow \infty} \frac{\alpha}{\alpha + \beta} = 1 \quad (5-6)$$

Teniendo en cuenta los límites anteriores, es evidente que $\mu \in (0, 1)$. Este resultado, también puede obtenerse con algo de análisis, pues dado el espacio numérico donde tiene sentido la variable aleatoria, lo lógico es que su valor esperado se mantenga en el mismo espacio. De la misma manera, es necesario realizar este proceso para σ^2 . Por ende, se desarrolló los límites igual que en el caso anterior, con el objetivo de establecer el rango de este indicador de dispersión.

Entonces, preliminarmente se debe considerar que:

$$\sigma^2 = \frac{\alpha}{(\alpha + \beta)} \frac{\beta}{(\alpha + \beta)(\alpha + \beta + 1)} = \mu \omega_{(\alpha, \beta)} \quad (5-7)$$

Por ello, antes de hallar los límites para σ^2 , se tiene que los mismos para la función $\omega_{(\alpha, \beta)}$ están determinados por:

$$\lim_{\beta \rightarrow \infty} \omega_{(\alpha, \beta)} = \lim_{\beta \rightarrow \infty} \frac{\beta}{(\alpha + \beta)(\alpha + \beta + 1)} = 0 \quad (5-8)$$

$$\lim_{\beta \rightarrow 0} \omega_{(\alpha, \beta)} = \lim_{\beta \rightarrow 0} \frac{\beta}{(\alpha + \beta)(\alpha + \beta + 1)} = 0 \quad (5-9)$$

$$\lim_{\alpha \rightarrow \infty} \omega_{(\alpha, \beta)} = \lim_{\alpha \rightarrow \infty} \frac{\beta}{(\alpha + \beta)(\alpha + \beta + 1)} = 0 \quad (5-10)$$

$$\lim_{\alpha \rightarrow 0} \omega_{(\alpha, \beta)} = \lim_{\alpha \rightarrow 0} \frac{\beta}{(\alpha + \beta)(\alpha + \beta + 1)} = \frac{\beta}{\beta(\beta + 1)} \quad (5-11)$$

Así, por las propiedades del límite y la expresión planteada en la ecuación (5-7), se puede deducir que:

$$\lim_{\beta \rightarrow \infty} \sigma^2 = \lim_{\beta \rightarrow \infty} \mu \omega_{(\alpha, \beta)} = \lim_{\beta \rightarrow \infty} \mu \lim_{\beta \rightarrow \infty} \omega_{(\alpha, \beta)} = 0(0) = 0 \quad (5-12)$$

Análogamente, se puede establecer,

$$\lim_{\beta \rightarrow 0} \sigma^2 = \lim_{\beta \rightarrow 0} \mu \omega_{(\alpha, \beta)} = 1(0) = 0 \quad (5-13)$$

$$\lim_{\alpha \rightarrow \infty} \sigma^2 = \lim_{\alpha \rightarrow \infty} \mu \omega_{(\alpha, \beta)} = 1(0) = 0 \quad (5-14)$$

$$\lim_{\alpha \rightarrow 0} \sigma^2 = \lim_{\alpha \rightarrow 0} \mu \omega_{(\alpha, \beta)} = 0 \left(\frac{\beta}{\beta(\beta + 1)} \right) = 0 \quad (5-15)$$

Apoyandose en el resultado anterior, si la intención es definir un rango para σ^2 , analizar los límites no proporciona suficiente información. Dado que queda en evidencia, que la variabilidad de una variable aleatoria con distribución beta, se aproxima a cero cuando alguno de los parámetros de forma tiende a cero. Así mismo, se aproxima a cero cuando los mencionados parámetros tienden a infinito.

En búsqueda de alternativas, se utilizaron herramientas que permitan detallar gráficamente el rango en cuestión, como se muestra en la Figura **5-1**, que expone el comportamiento de σ^2 en el espacio, a medida que varían los diferentes valores de α y β . En ésta se puede observar que el máximo valor que este indicador de dispersión puede tomar, está muy cercano a 0,25. Además, de manera empírica se calculan los posibles valores que σ^2 puede tomar, asignando valores desde 0 hasta 2000 a α , para cada valor de β en ese mismo intervalo. Lo anterior, se ve reflejado en la Figura **5-2**, que muestra los valores de σ^2 frente a β , y cada curva representando cada valor de α . Donde se refuerza la idea de que ésta siempre es menor que 0,25, y concuerda con lo inferido en el cálculo de sus límites. Pues sin importar el valor de α , a medida que los valores del parámetro β se acercan a cero o incrementan después de 2, se observa que σ^2 siempre tiende a cero. En consecuencia, es razonable pensar que la variabilidad de la distribución beta siempre está dentro del intervalo (0, 0,25).

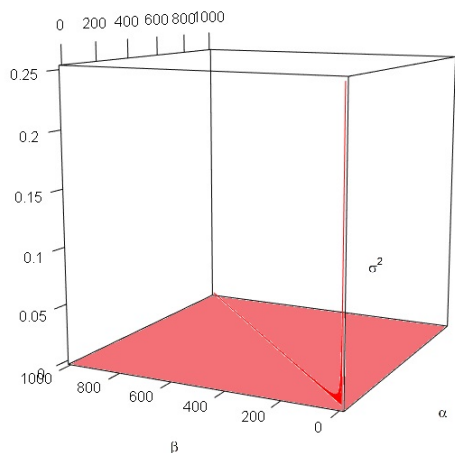


Figura 5-1: Comportamiento de σ^2 vs Valores fijos de (α, β)

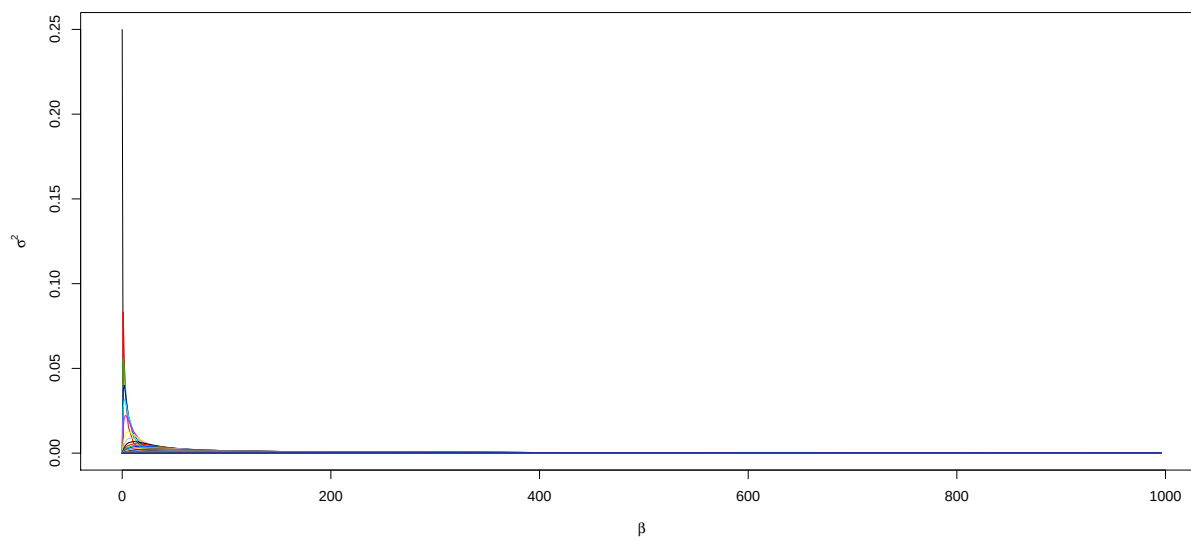


Figura 5-2: Comportamiento empírico de σ^2

Por otra parte, como μ y σ^2 están en términos de los parámetros de forma, lo que se pretende es plantear una distribución a priori conjunta $f(\mu, \sigma^2)$, de forma tal, que empleando las transformaciones bivariadas de distribuciones de variables aleatorias, se pueda derivar en una a priori conjunta de la manera $f(\alpha, \beta)$, y finalmente marginalizar $f(\alpha)$ y $f(\beta)$. En este

contexto, es necesario conocer si la transformación es uno a uno. Hasta el momento, se sabe que valores positivos de α y β establecen que $\mu \in (0, 1)$ y $\sigma^2 \in (0, 0,25)$ en (5-2).

Ahora, se requiere conocer si estos últimos derivan en valores positivos para los parámetros de forma en la transformación inversa, mostrada en las ecuaciones (5-16) y (5-17).

$$\alpha = \mu \left[\frac{\mu(1-\mu)}{\sigma^2} - 1 \right] \quad (5-16) \quad \beta = (1-\mu) \left[\frac{\mu(1-\mu)}{\sigma^2} - 1 \right] \quad (5-17)$$

Así pues, analíticamente se puede concluir que para que α y β pertenezcan al intervalo $(0, \infty)$, que es la restricción de la función de densidad beta, se debe cumplir que $\frac{\mu(1-\mu)}{\sigma^2} - 1 > 0$, por ende:

$$\sigma^2 < \mu(1-\mu) \quad (5-18)$$

Esto es, para que α y β sean positivos, μ puede tomar cualquier valor dentro del intervalo $(0, 1)$. Sin embargo, pese a que es correcto que σ^2 siempre está en $(0, 0,25)$, ésta no puede tomar cualquier valor libremente en ese intervalo, pues si no se cumple lo expresado en la ecuación (5-18), α y β pueden presentar valores negativos, es decir que exactamente $\sigma^2 \in (0, \mu(1-\mu))$. Lo anterior, deja en evidencia que σ^2 depende de los valores que pueda tomar μ , lo cual se explica gráficamente en la Figura **5-3**, que representa el comportamiento de la varianza frente a la media, en sus respectivos intervalos. Allí, el área sombreada representa el rango donde la varianza de la distribución beta tiene sentido, que es bajo la curva de la parábola $\mu(1-\mu)$, y es donde la transformación en cuestión presenta la característica de ser uno a uno y cumplir con la restricción de una densidad beta, donde los parámetros de forma siempre deben ser positivos.

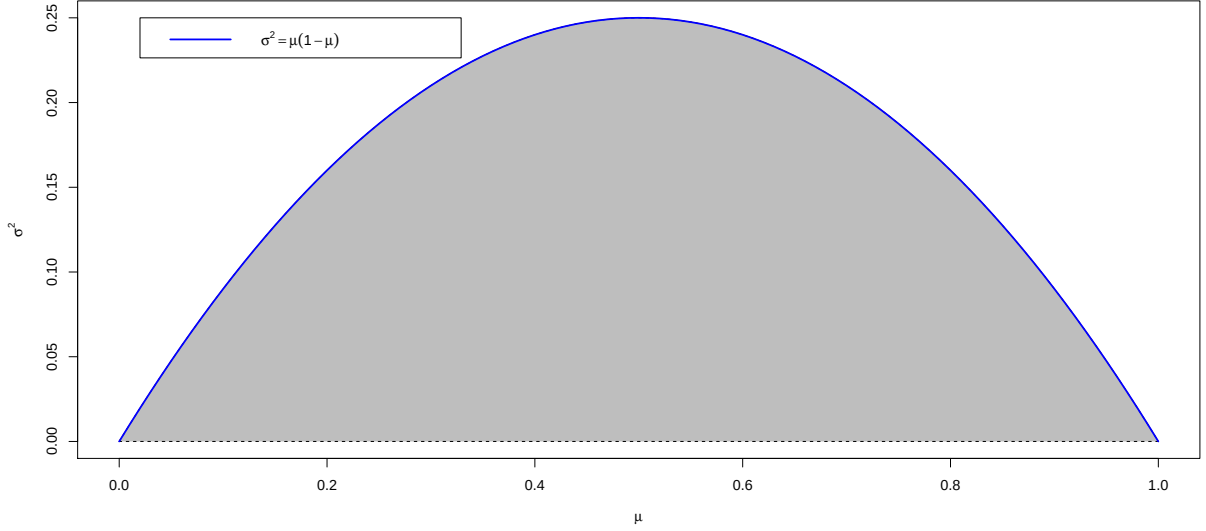


Figura 5-3: Gráfica de σ^2 dado los valores de μ

Basandose en los resultados anteriores, es coherente considerar una distribución a priori para σ^2 condicional en μ , de la manera $f(\sigma^2|\mu)$. También, una distribución a priori para μ , $f(\mu)$, tal que se puede expresar una distribución a priori conjunta, teniendo en cuenta la definición de probabilidad condicional, de la forma:

$$f(\mu, \sigma^2) = f(\mu)f(\sigma^2|\mu) \quad (5-19)$$

Ya que es una función de densidad flexible para modelar variables aleatorias continuas, restringidas en el intervalo $(0, 1)$, y que la integral de su función de densidad es de forma cerrada, se emplea una distribución a priori Kumaraswamy con hiperparámetros a_1 y b_1 , para modelar el conocimiento que se tiene de μ previamente. Entonces:

$\mu \sim \text{Kumaraswamy}(a_1, b_1)$, con función de densidad,

$$f(\mu) = a_1 b_1 \mu^{a_1-1} (1 - \mu^{a_1})^{b_1-1} \quad ; 0 < \mu < 1; \quad a_1 > 0, \quad b_1 > 0 \quad (5-20)$$

Del mismo modo, se le asigna la misma importancia, o la misma probabilidad de ocurrencia a cualquier valor de σ^2 dentro de la región sombreada en la Figura 5-3. Por ello, se establece una a priori condicional uniforme no informativa, expresada como sigue:

$\sigma^2|\mu \sim Uniforme(0, \mu(1 - \mu))$, así,

$$f(\sigma^2|\mu) = \frac{1}{\mu(1 - \mu)} \quad ; 0 < \sigma^2 < \mu(1 - \mu) \quad (5-21)$$

Por lo tanto, ya se tiene la estructura suficiente para establecer una a priori conjunta para la media y la varianza de la distribución Beta, como se muestra en la ecuación (5-19). En consecuencia, se tiene que:

$$f(\mu, \sigma^2) = \frac{a_1 b_1}{\mu(1 - \mu)} \mu^{a_1-1} (1 - \mu^{a_1})^{b_1-1} \quad (5-22)$$

Este escenario, plantea que el proceso de elicitación solo queda dirigido a encontrar los valores de los hiperparámetros a_1 y b_1 . Lo cual sugiere que bajo este enfoque, sólo se debe tener conocimiento previo del valor esperado de la variable aleatoria de interés, lo que facilita el proceso de elicitación. Ya que para los expertos, puede ser más simple entender, interpretar, y/o suministrar información sobre el valor promedio de algún hecho o circunstancia del tema que incumbe la investigación.

Una vez obtenida la forma de la distribución de densidad previa conjunta para μ y σ^2 , como se muestra en la ecuación (5-22), es necesario encontrar la forma de la distribución de densidad previa conjunta para los parámetros α y β de la forma en que se observa en la ecuación (5-23), y posteriormente obtener las marginales de esta. Para ello, como se mencionó anteriormente, se empleó la metodología de las transformaciones bivariadas de variables aleatorias.

$$f(\alpha, \beta) = f(\mu, \sigma^2)|J| \quad (5-23)$$

Esta metodología hace uso de una cantidad llamada la transformación del jacobiano, denotado por $|J|$, el cual es el determinante de la matriz de derivadas parciales de μ y de σ^2 respecto a α y β , como se muestra a continuación.

$$J = \begin{pmatrix} \frac{\partial \mu}{\partial \alpha} & \frac{\partial \mu}{\partial \beta} \\ \frac{\partial \sigma^2}{\partial \alpha} & \frac{\partial \sigma^2}{\partial \beta} \end{pmatrix}$$

Entonces, al realizar el cálculo de las derivadas parciales de μ respecto a cada parámetro de forma, se obtuvieron las siguientes expresiones:

$$\frac{\partial \mu}{\partial \alpha} = \frac{(\alpha + \beta) - \alpha}{(\alpha + \beta)^2} = \frac{\beta}{(\alpha + \beta)^2} \quad (5-24)$$

$$\frac{\partial \mu}{\partial \beta} = \frac{-\alpha}{(\alpha + \beta)^2} \quad (5-25)$$

Análogamente, se realiza el procedimiento para obtener las expresiones de las derivadas parciales de σ^2 respecto a α y β , como se muestra en las ecuaciones (5-26) y (5-27) respectivamente.

$$\begin{aligned} \frac{\partial \sigma^2}{\partial \alpha} &= \frac{\beta[(\alpha + \beta)^2(\alpha + \beta + 1)] - \alpha\beta[2(\alpha + \beta)(\alpha + \beta + 1) + (\alpha + \beta)^2]}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \\ &= \frac{\beta[\alpha^3 + \beta^3 + \alpha^2 + \beta^2 + 3\alpha^2\beta + 3\alpha\beta^2 + 2\alpha\beta] - \alpha\beta[3\alpha^2 + 6\alpha\beta + 2\alpha + 2\beta + 3\beta^2]}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \\ &= \frac{\beta\alpha^3 + \beta^4 + \beta\alpha^2 + \beta^3 + 3\alpha^2\beta^2 + 3\alpha\beta^3 + 2\alpha\beta^2 - 3\alpha^3\beta - 6\alpha^2\beta^2 - 2\alpha^2\beta - 2\alpha\beta^2 - 3\alpha\beta^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \end{aligned}$$

$$\frac{\partial \sigma^2}{\partial \alpha} = \frac{\beta^4 - 2\alpha^3\beta - \alpha^2\beta - 3\alpha^2\beta^2 + \beta^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \quad (5-26)$$

$$\begin{aligned} \frac{\partial \sigma^2}{\partial \beta} &= \frac{\alpha[(\alpha + \beta)^2(\alpha + \beta + 1)] - \alpha\beta[2(\alpha + \beta)(\alpha + \beta + 1) + (\alpha + \beta)^2]}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \\ &= \frac{\beta^3\alpha + \alpha^4 + \beta^2\alpha + \alpha^3 + 3\alpha^3\beta + 3\alpha^2\beta^2 + 2\alpha^2\beta - 3\alpha^3\beta - 6\alpha^2\beta^2 - 2\alpha^2\beta - 2\alpha\beta^2 - 3\alpha\beta^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \end{aligned}$$

$$\frac{\partial \sigma^2}{\partial \beta} = \frac{\alpha^4 - 2\alpha\beta^3 - \alpha\beta^2 - 3\alpha^2\beta^2 + \alpha^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \quad (5-27)$$

Ahora, se procede a calcular el determinante de dicha matriz. Finalmente, el Jacobiano está definido de la siguiente manera:

$$|J| = \frac{\partial \mu}{\partial \alpha} \frac{\partial \sigma^2}{\partial \beta} - \frac{\partial \mu}{\partial \beta} \frac{\partial \sigma^2}{\partial \alpha} \quad (5-28)$$

Reemplazando las expresiones (5-24), (5-25), (5-26), (5-27) en (5-28), este queda determinado por:

$$\begin{aligned}
|J| &= \left(\frac{\beta}{(\alpha + \beta)^2} \right) \frac{\alpha^4 - 2\alpha\beta^3 - \alpha\beta^2 - 3\alpha^2\beta^2 + \alpha^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} - \left(\frac{-\alpha}{(\alpha + \beta)^2} \right) \frac{\beta^4 - 2\alpha^3\beta - \alpha^2\beta - 3\alpha^2\beta^2 + \beta^3}{(\alpha + \beta)^4(\alpha + \beta + 1)^2} \\
&= \frac{\alpha^4\beta - 2\alpha\beta^4 + \alpha^3\beta - \alpha\beta^3 - 3\alpha^2\beta^3 + \alpha\beta^4 - 2\alpha^4\beta - \alpha^3\beta - 3\alpha^3\beta^2 + \beta^3\alpha}{(\alpha + \beta)^6(\alpha + \beta + 1)^2} \\
&= \frac{\alpha\beta(\alpha^3 - 2\beta^3 + \alpha^2 - \beta^2 - 3\alpha\beta^2 + \beta^3 - 2\alpha^3 - \alpha^2 - 3\alpha^2\beta + \beta^2)}{(\alpha + \beta)^6(\alpha + \beta + 1)^2} \\
&= \frac{\alpha\beta(-\alpha^3 - \beta^3 - 3\alpha\beta^2 - 3\alpha^2\beta)}{(\alpha + \beta)^6(\alpha + \beta + 1)^2} \\
&= \frac{\alpha\beta(-\alpha - \beta)^3}{(\alpha + \beta)^6(\alpha + \beta + 1)^2} = \frac{-\alpha\beta(\alpha + \beta)^3}{(\alpha + \beta)^6(\alpha + \beta + 1)^2}
\end{aligned}$$

$$|J| = \left| \frac{-\alpha\beta}{(\alpha + \beta)^3(\alpha + \beta + 1)^2} \right| \quad (5-29)$$

Por lo tanto, al reemplazar (5-22) y (5-29) en (5-23), se obtiene la forma de la distribución a priori conjunta $f(\alpha, \beta)$ como se muestra a continuación.

$$f(\alpha, \beta) = \frac{a_1 b_1}{\left(\frac{\alpha}{\alpha + \beta}\right) \left(1 - \frac{\alpha}{\alpha + \beta}\right)} \cdot \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1 - 1} \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1 - 1} \left[\frac{\alpha\beta}{(\alpha + \beta)^3(\alpha + \beta + 1)^2} \right] \quad (5-30)$$

Ahora bien, teniendo en cuenta el objetivo principal del presente documento, es necesario marginalizar la expresión anterior, en búsqueda de una expresión para la distribuciones a priori marginales de α y β respectivamente.

De ahí que, es necesario integrar la distribución a priori conjunta mostrada en (5-30) con respecta a β , con el propósito de obtener la distribución a priori de α . De éste modo, se tiene que:

$$f(\alpha) = \int_0^\infty f(\alpha, \beta) d\beta \quad (5-31)$$

Para obtener una solución la expresión anterior, se lleva a cabo la integración por partes. Entonces:

$$f(\alpha) = \left[uv - \int v du \right]_0^\infty \quad (5-32)$$

En primera instancia dv y u están determinadas respectivamente por:

$$dv = \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1 - 1} \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1 - 1} \left(\frac{\alpha}{(\alpha + \beta)^2}\right) d\beta \quad (5-33)$$

$$\begin{aligned}
u &= \frac{1}{\left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\beta}{\alpha+\beta}\right)} \left[\frac{\beta}{(\alpha+\beta)(\alpha+\beta+1)^2} \right] \\
&= \frac{\alpha+\beta}{\alpha(\alpha+\beta+1)^2}
\end{aligned} \tag{5-34}$$

Lo cual, sugiere que v y du se resuelven de la siguiente manera:

$$\begin{aligned}
v &= \int dv = -\left[1 - \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1}\right] \\
&= \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} - 1
\end{aligned} \tag{5-35}$$

$$du = -\left[\frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3}\right]d\beta \tag{5-36}$$

Por ende, ya que se conoce la forma teórica de u y v , para solucionar la ecuación (5-32) solo es necesario resolver la siguiente integral.

$$\begin{aligned}
\int vdu &= - \int \left(\left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} - 1 \right) \frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3} d\beta \\
&= - \int \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} \frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3} d\beta + \int \frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3} d\beta
\end{aligned} \tag{5-37}$$

Ahora, considerando que si se escribe

$$\begin{aligned}
\int vdu &= \Omega - \int du \\
&= \Omega - u
\end{aligned} \tag{5-38}$$

Donde,

$$\Omega = - \int \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} \frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3} d\beta \tag{5-39}$$

Por consiguiente, reemplazando (5-38) en (5-32), se tiene que la distribución a priori de α , es de la forma:

$$\begin{aligned}
f(\alpha) &= [uv - (\Omega - u)]_0^\infty \\
&= [u(1+v)]_0^\infty - \Omega]_0^\infty
\end{aligned} \tag{5-40}$$

Remplazando los valores de u , v y Ω en términos de los parámetros de la distribución Beta α y β ,

$$f(\alpha) = \frac{\alpha+\beta}{\alpha(\alpha+\beta+1)^2} \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} \Big|_0^\infty + \int_0^\infty \left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1} \frac{\alpha+\beta-1}{\alpha(\alpha+\beta+1)^3} d\beta \tag{5-41}$$

Finalmente, se obtiene:

$$f(\alpha|a_1, b_1) = \int_0^\infty \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1} \frac{\alpha + \beta - 1}{\alpha(\alpha + \beta + 1)^3} d\beta \quad (5-42)$$

Ésta última, es una integral que no tiene forma cerrada. Ello conlleva, a que se tenga que utilizar un método numérico para aproximar la densidad en distintos valores del parámetro α , cuando se han elicitado los términos de los hiperparámetros a_1 y b_1 . En este contexto, a continuación se expone un algoritmo para aproximar dicha densidad empíricamente, empleando la regla de los trapecios. Previamente, por practicidad se define la función de densidad a priori como: $f(\alpha|a_1, b_1) = \int_0^\infty g(\alpha, \beta) d\beta$. Donde $g(\alpha, \beta) = \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1} \frac{\alpha + \beta - 1}{\alpha(\alpha + \beta + 1)^3}$.

Algoritmo de aproximación

1. Obtener los valores de los hiperparámetros a_1 y b_1 , construyendo la a priori de μ .
2. Definir valores para el parámetro de forma α , desde un valor muy cercano a cero α_0 hasta uno muy grande α_M . En este último, de manera empírica se evidenció que generalmente funciona para valores mayores o iguales a 100.
3. Determinar el número de trapecios n y el número máximo donde se definirá la integral b (En este caso debe ser un número muy grande, pues homologa el ∞ de la integral objetivo). Además, en este caso $a=0$. Esto significa, que se está integrando la marginal de α , en el espacio definido para β en (a, b) .
4. Inicializar un bucle que recorrerá la posición de los valores de α previamente definidos en (α_0, α_M) , el cual será recorrido por el contador j , desde $j = 1$.
5. Inicializar un bucle que recorrerá la posición de los trapecios previamente definidos en (a, b) , el cual será recorrido por el contador i , desde $i = 1$.
6. Especificar un valor para el parámetro β , el cual es igual a la posición i -ésima de los trapecios.
7. Evaluar la función $g(\alpha, \beta)$. Donde el valor de α es igual a la posición j -ésima del intervalo (α_0, α_M) .
8. Si i es igual a 1 o b , multiplicar la expresión anterior por la constante $\frac{b}{2n}$. De lo contrario, multiplicarla por $\frac{b}{n}$.
9. Almacenar el resultado preliminar en la posición i -ésima de un vector de densidades (Este vector se debe crear previamente, y tiene la función de almacenar las densidades encontradas en el paso anterior).
10. Aumentar el contador tal que $i = i + 1$, y repetir los pasos desde el 6 hasta el 9, b veces.
11. Sumar todas las densidades almacenadas en el vector de densidades, y almacenarlas en la posición j -ésima de un vector de áreas (Esta será la aproximación por la regla de los trapecios para el valor j -ésimo de α en el intervalo (α_0, α_M))

12. Aumentar el contador tal que $j = j + 1$, y repetir los pasos desde el 5 al 11, tantas veces como valores de α se definan en el intervalo (α_0, α_M) .

Finalmente, el vector de áreas obtenido, es una aproximación empleando la metodología de la regla de los trapecios, para valores definidos de α dentro de su espacio paramétrico, de la distribución a priori marginal de dicho parámetro.

En el siguiente capítulo, se evidenciará que al ejecutar el algoritmo mostrado preliminarmente, para valores dados de α , la integral (5-42) tiende a cero rápidamente.

Luego, de forma análoga al procedimiento realizado anteriormente para obtener la distribución a priori para el parámetro α , se integró la distribución a priori conjunta $f(\alpha, \beta)$. Sin embargo, esta vez se tuvo en cuenta que la integración se desarrolló respecto al parámetro de forma α , como se puede ver expresado en (5-43).

$$f(\beta) = \int_0^\infty f(\alpha, \beta) d\alpha \quad (5-43)$$

Para la solución de esta integral, se hace necesario aplicar nuevamente el método de integración por partes. Por lo tanto, se reescribió $f(\beta)$ de la manera que se puede ver expresado en (5-44), y posteriormente se definieron sus respectivos factores.

$$f(\beta) = uv - \int v du \quad (5-44)$$

Donde dv y u son las ecuaciones (5-45) y (5-46) respectivamente.

$$dv = \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1 - 1} \left(1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right)^{b_1 - 1} \left(\frac{\beta}{(\alpha + \beta)^2} \right) d\alpha \quad (5-45)$$

$$u = \frac{1}{\left(\frac{\alpha}{\alpha + \beta} \right) \left(\frac{\beta}{\alpha + \beta} \right)} \left[\frac{\alpha}{(\alpha + \beta)(\alpha + \beta + 1)^2} \right]$$

$$u = \frac{\alpha + \beta}{\beta(\alpha + \beta + 1)^2} \quad (5-46)$$

Acto seguido, se realizó la derivada parcial de u respecto al parámetro α (5-47), y a su vez, se desarrolló la integral de dv con la intención de obtener la expresión de v (5-48). De esta manera, quedaron definidos formalmente los respectivos factores que son necesarios para hacer uso del método de integración por partes, y así, hallar la distribución a priori para el parámetro de forma β .

$$\begin{aligned}
\frac{du}{d\alpha} &= \frac{\beta(\alpha + \beta + 1)^2 - (\alpha + \beta)2\beta(\alpha + \beta + 1)}{\beta^2(\alpha + \beta + 1)^4} \\
&= \frac{\beta(\alpha + \beta + 1)[(\alpha + \beta + 1) - 2(\alpha + \beta)]}{\beta^2(\alpha + \beta + 1)^4} \\
&= \frac{1 - \alpha - \beta}{\beta(\alpha + \beta + 1)^3}
\end{aligned}$$

$$\frac{du}{d\alpha} = -\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \quad (5-47)$$

$$v = \int dv = 1 - \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \quad (5-48)$$

Entonces, al reemplazar (5-46), (5-47) y (5-48) en (5-44), se tiene que:

$$f(\beta) = uv \Big|_0^\infty - \int_0^\infty v du$$

Por practicidad, se decidió trabajar primero la parte que contiene la integral, con la intención de simplificar la expresión. Entonces se tiene que:

$$\begin{aligned}
\int v du &= \int \left[1 - \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \right] \left[-\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha \\
&= \int \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \left[\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha - \int \left[\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha \\
&= \Omega + u
\end{aligned}$$

Donde

$$\Omega = \int \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \left[\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha$$

Y, se puede observar que la otra parte de ésta, corresponde a la integral de la ecuación (5-47), cuya forma hace alusión a lo que se había definido previamente como u .

Por lo tanto, $f(\beta)$ se puede expresar de la siguiente forma, la cual posteriormente será evaluada en los respectivos límites de integración $(0, \infty)$

$$\begin{aligned}
f(\beta) &= uv \big|_0^\infty - u \big|_0^\infty - \Omega \\
&= u[v - 1] \big|_0^\infty - \Omega \\
&= \left[\left[\frac{\alpha + \beta}{\beta(\alpha + \beta - 1)^2} \right] \left[\left[1 - \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \right] - 1 \right] \right]_0^\infty - \Omega \\
&= \left[\left[\frac{\alpha + \beta}{\beta(\alpha + \beta - 1)^2} \right] \left[\left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \right] \right]_0^\infty - \Omega \\
&= \frac{\beta}{\beta(\beta + 1)^2} - \Omega
\end{aligned}$$

Finalmente, la forma obtenida de la distribución a priori del parámetro de forma β , se puede observar a continuación en la ecuación (5-49)

$$f(\beta) = \frac{1}{(\beta + 1)^2} - \int_0^\infty \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \left[\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha \quad (5-49)$$

Sin embargo, para resolver la integral que se encuentra dentro de la ecuación (5-49) se recurrió de nuevo al método de aproximación de la regla de los trapecios. Y, análogamente al procedimiento realizado para obtener el valor de la integral de la distribución marginal $f(\alpha)$, se dio solución a dicha integral, cuyo resultado se muestra a continuación en la ecuación (5-50).

$$\int_0^\infty \left[1 - \left(\frac{\alpha}{\alpha + \beta} \right)^{a_1} \right]^{b_1} \left[\frac{\alpha + \beta - 1}{\beta(\alpha + \beta - 1)^3} \right] d\alpha \approx 0 \quad (5-50)$$

Lo cual sugiere, que esta parte de la densidad marginal de β es un error despreciable, un residuo que no altera la densidad de este parámetro de forma. Por lo tanto, la forma de $f(\beta)$ quedó reducida como se puede observar en la ecuación (5-51).

$$f(\beta) = \frac{1}{(\beta + 1)^2} \quad (5-51)$$

Ahora bien, para probar si efectivamente la distribución previa marginal del parámetro de forma β es una densidad, se integró $f(\beta)$ de la siguiente manera.

$$\int_0^\infty f(\beta) d\beta = \left[-\frac{1}{\beta + 1} \right]_0^\infty = 1 \quad (5-52)$$

Dado este resultado, se puede decir que el área total encerrada bajo la curva es igual a 1, tal como se muestra en la ecuación (5-52), se comprueba que la expresión que se obtuvo como distribución previa marginal de β es una distribución de densidad. Además, la forma acumulada de ésta se muestra a continuación en la ecuación (5-53)

$$F(\beta) = 1 - \frac{1}{\beta + 1} \quad (5-53)$$

Ahora bien, al poseer una forma analítica para la distribución acumulada a priori de β , se empleó el método de la transformada inversa para obtener números aleatorios de ésta distribución. Por ello, se establece que si se tiene una variable aleatoria u , de tal manera que $u \sim \text{Uniforme}(0,1)$, entonces $F(\beta) = u$. Lo que implica que $\beta = F^{-1}(u)$, así:

$$u = 1 - \frac{1}{\beta + 1} \quad (5-54)$$

Lo expresión anterior, conlleva a que finalmente sea:

$$\beta = \frac{u}{1 - u} \quad (5-55)$$

De este modo, los números generados a través de la ecuación (5-55), pertenecen a los que pueden ser originados a partir de la distribución a priori marginal de β .

Adicionalmente, si se desea obtener la forma de las distribuciones condicionales $f(\alpha|\beta)$ y $f(\beta|\alpha)$, se debe tener en cuenta que.

$$f(\alpha|\beta) = \frac{f(\alpha, \beta)}{f(\beta)} \quad ; \quad f(\beta|\alpha) = \frac{f(\alpha, \beta)}{f(\alpha)} \quad (5-56)$$

Entonces, al dividir la función de distribución conjunta (5-30) sobre de la función de distribución marginal (5-51), se obtiene la expresión de la distribución condicional $f(\alpha|\beta)$ como se puede observar a continuación en la ecuación (5-57). Del mismo modo, se divide la función de distribución conjunta $f(\alpha, \beta)$ sobre la distribución marginal (5-42) para obtener $f(\beta|\alpha)$, como se observa en 5-58.

$$f(\alpha|\beta) = \frac{a_1 b_1 (\beta + 1)^2}{\left(\frac{\alpha}{\alpha + \beta}\right) \left(1 - \frac{\alpha}{\alpha + \beta}\right)} \cdot \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1 - 1} \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1 - 1} \left[\frac{\alpha \beta}{(\alpha + \beta)^3 (\alpha + \beta + 1)^2} \right] \quad (5-57)$$

$$f(\beta|\alpha) = \frac{\frac{a_1 b_1}{\left(\frac{\alpha}{\alpha + \beta}\right) \left(1 - \frac{\alpha}{\alpha + \beta}\right)} \cdot \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1 - 1} \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1 - 1} \left[\frac{\alpha \beta}{(\alpha + \beta)^3 (\alpha + \beta + 1)^2} \right]}{\int_0^\infty \left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^{a_1}\right)^{b_1} \frac{\alpha + \beta - 1}{\alpha(\alpha + \beta + 1)^3} d\beta} \quad (5-58)$$

Estas distribuciones condicionales, se hallan con la intención de encontrar expresiones adicionales al momento de plantear una distribución posterior conjunta para los parámetros de la distribución Beta. Además, de que estas sirvan como insumos para la realización de futuros trabajos o investigaciones, respecto al tratamiento Bayesiano que se le pueda dar a la distribución Beta.

Ahora bien, basándose en la forma de la expresión del teorema de Bayes, cuya ecuación se puede observar en (4-19), la forma de obtener esta distribución posterior se muestra a continuación en la

ecuación (5-59).

$$f(\alpha, \beta|x) = cL(\alpha, \beta|x)f(\alpha, \beta) = cL(\alpha, \beta|x)f(\alpha|\beta)f(\beta) = cL(\alpha, \beta|x)f(\beta|\alpha)f(\alpha) \quad (5-59)$$

Donde la expresión que se observa en (5-60), corresponde a la verosimilitud de la distribución Beta, y c es la constante de normalización, que viene dada por la ecuación (5-61)

$$L(\alpha, \beta|x) = \prod_{i=1}^n \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1 - x_i)^{\beta-1} \quad (5-60)$$

$$c = \frac{1}{\int \int L(\alpha, \beta|x) f(\alpha, \beta) d\beta d\alpha} \quad (5-61)$$

De este modo, para trabajos posteriores respecto a la estimación bayesiana, se puede partir de cualquiera de las formas para obtener la distribución posterior $f(\alpha, \beta|x)$ planteadas en la ecuación (5-59), que explícitamente está expresado de la manera en que se muestra en la ecuación (5-62)

$$f(\alpha, \beta|x) = \frac{\left[\prod_{i=1}^n \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1-x_i)^{\beta-1} \right] \frac{a_1 b_1}{\left(\frac{\alpha}{\alpha+\beta}\right) \left(1-\frac{\alpha}{\alpha+\beta}\right)} \cdot \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1-1} \left(1-\left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1-1} \left[\frac{\alpha\beta}{(\alpha+\beta)^3 (\alpha+\beta+1)^2} \right]}{\int_0^\infty \int_0^\infty \left[\prod_{i=1}^n \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1-x_i)^{\beta-1} \right] \frac{a_1 b_1}{\left(\frac{\alpha}{\alpha+\beta}\right) \left(1-\frac{\alpha}{\alpha+\beta}\right)} \cdot \left(\frac{\alpha}{\alpha+\beta}\right)^{a_1-1} \left(1-\left(\frac{\alpha}{\alpha+\beta}\right)^{a_1}\right)^{b_1-1} \left[\frac{\alpha\beta}{(\alpha+\beta)^3 (\alpha+\beta+1)^2} \right] d\beta d\alpha} \quad (5-62)$$

Partiendo de lo anterior se deben de hacer los cálculos matemáticos correspondientes y evaluar si estudiar la distribución posterior con la metodología expuesta hasta aquí es pertinente, o si existe la posibilidad de trabajar conjuntamente con otras metodologías estadísticas o matemáticas, para finalmente encontrar las estimaciones Bayesianas.

6 Resultados

6.1. Distribución A Priori para α

Creando un escenario de simulación, se fijaron los valores de los parámetros de forma de la distribución Beta, tal que $\alpha = 0.15$ y $\beta = 1$. Lo anterior, implica que $\mu = 0.1304$.

Del mismo modo, se simuló el proceso de elicitación, donde se supone que entrevistando un experto se estableció un valor a priori para $\mu^* = 0,2104$, además especificando un valor arbitrario para $b_1 = 3$ y remplazando estos en la ecuación 4-18, se obtiene un valor para el hiperparámetro $a_1 = 1.0127$. En este contexto, la Figura 6-1 muestra la curva de densidad de la distribución a priori estimada, tal que $\mu \sim Kuma(a_1 = 1.0127, b_1 = 3)$.

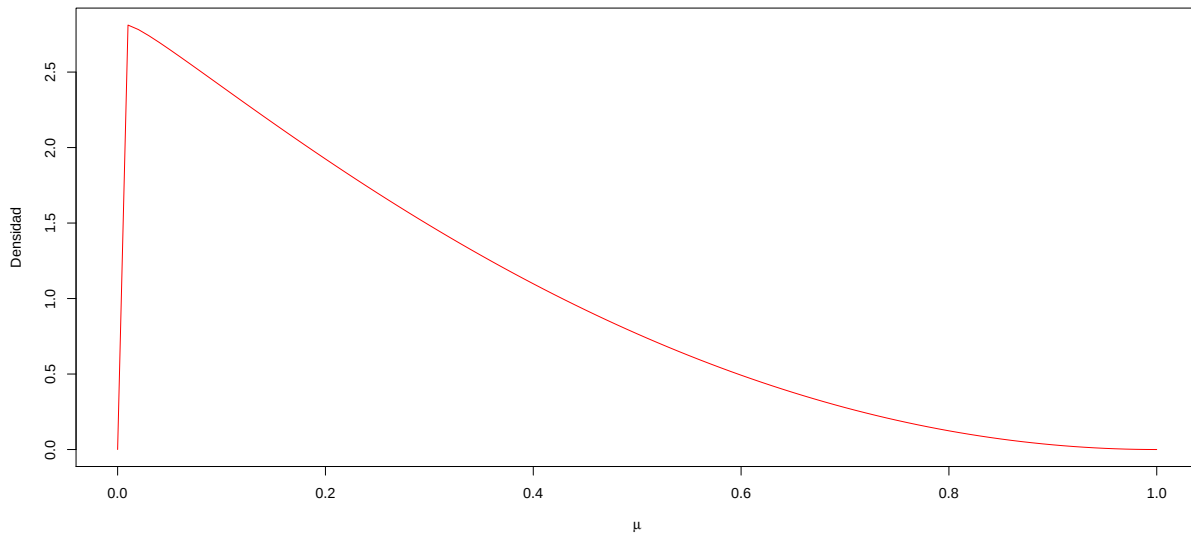


Figura 6-1: Curva de densidad de la distribución a priori de μ

A continuación, en la Figura 6-2 se presenta el resultado obtenido a partir del algoritmo de aproximación de la función de densidad expresada en la ecuación (5-42), el cual fue planteado en la sección anterior.

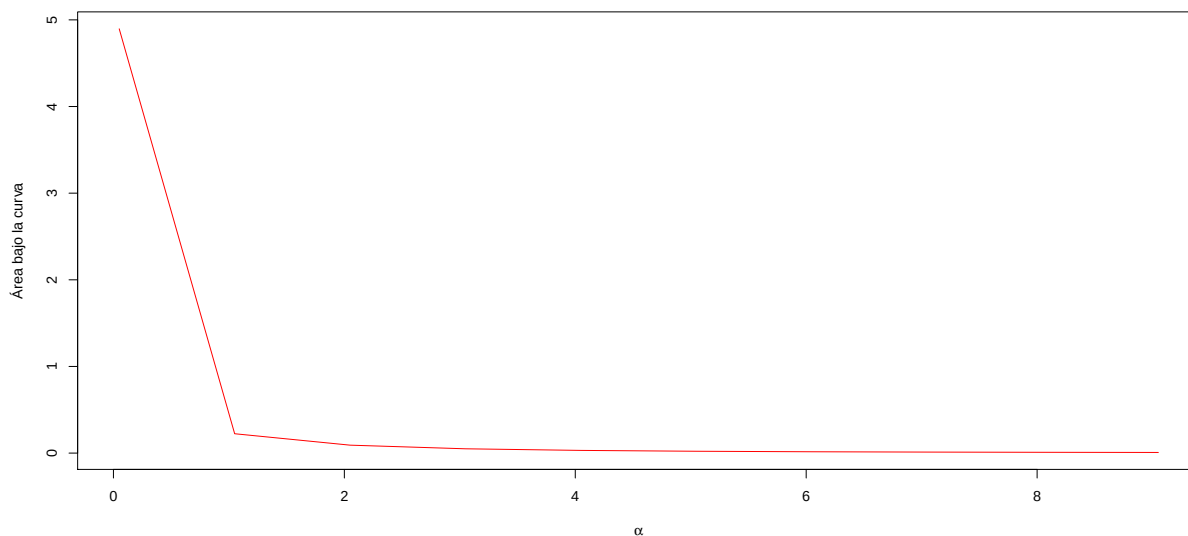


Figura 6-2: Curva de densidad de la distribución a priori aproximada con el algoritmo

Se puede evidenciar que la densidad aproximada, teniendo en cuenta los valores específicos de α , tiende a cero rápidamente. Particularmente, para cada valor del parámetro de forma α mayor o igual a 1, la suma de las densidad de la integral siempre es menor a 0.005. De la misma manera, el valor máximo de dicha suma se alcanza cuando α toma valores cercanos a cero. Específicamente, cuando $\alpha = 0.05$, la suma de las densidades bajo la aproximación de la regla de los trapecios es igual a 4,89.

En esa misma línea, la densidad mínima aproximada a través del algoritmo es de $1,7269e^{-07}$. Encontrándose estos valores tan pequeños generalmente cuando $\alpha > 10$. Además, se obtuvo que la suma de densidades promedio es de 0.0027.

Teniendo en cuenta lo anterior, se puede concluir, que partiendo de la función de densidad a priori conjunta para α y β , mostrada en la ecuación (5-30), es posible obtener una expresión para la distribución a priori marginal, donde es poco probable tener valores del parámetro de forma $\alpha > 1$. Sin embargo, ésta expresión no es de forma cerrada, por ende todos los resultados encontrados son aproximaciones. En otras palabras, empíricamente se puede pensar que tiene sentido utilizar esta distribución cuando en el proceso de elicitación, subjetivamente se plantea que α puede pertenecer al intervalo comprendido en $(0, 1)$.

Ahora bien, se acopló el algoritmo de aproximación expuesto en el capítulo anterior (estableciendo en el paso 2, un solo valor inicial para α , de tal forma que sea el valor que se halló en la elicitación de μ , y remplazando éste último en la ecuación (5-16)), en conjunto con el algoritmo Metropolis-Hasting, con la finalidad de generar números aleatorios que se aproximen a los

que pueden ser originados por la distribución a priori de α . Entonces, teniendo en cuenta el comportamiento observado a lo largo de la curva en la figura anterior, el espacio paramétrico y mediante una tarea de sensibilidad, se emplea una distribución propuesta independiente, de forma que $\phi \sim \exp(\lambda = 1)$. Por esto, $h(\phi) = e^{-\phi}$. Partiendo de lo anterior, a continuación se expone el algoritmo mediante el cual se puede generar una muestra aleatoria de la distribución a priori de α

Algoritmo de muestreo

1. Obtener los valores de los hiperparámetros a_1 y b_1 , construyendo la a priori de μ .
2. Con el valor de μ obtenido, generar un valor aleatorio para σ^2 , a partir de la distribución a priori condicional $Uniforme(0, \mu(1 - \mu))$
3. Definir un valor inicial para el parámetro de forma α^0 , de tal forma que este sea obtenido reemplazando los valores de μ y σ^2 definidos en los pasos anteriores, y reemplazando en $\alpha^0 = \mu \left[\frac{\mu(1-\mu)}{\sigma^2} - 1 \right]$.
4. Determinar el número de trapecios n y el número máximo donde se definirá la integral b (En este caso debe ser un número muy grande, pues homologa el ∞ de la integral objetivo). Además, en este caso $a=0$. Esto significa, que se está integrando la marginal de α , en el espacio definido para β en (a, b) .
5. Inicializar un bucle que definirá la cantidad de iteraciones que tendrá el algoritmo Metropolis-Hasting, el cual será recorrido por el contador j , desde $j = 1$.
6. Generar un valor aleatorio de ϕ , que será candidato para α , a partir de la distribución propuesta $\exp(\lambda = 1)$
7. Inicializar un bucle que recorrerá la posición de los trapecios previamente definidos en (a, b) , el cual será recorrido por el contador i , desde $i = 1$.
8. Especificar un valor para el parámetro β , el cual es igual a la posición i -ésima de los trapecios.
9. Evaluar las funciones $g(\alpha^{j-1}, \beta)$ y $g(\phi, \beta)$.
10. Si i es igual a 1 o b , multiplicar las expresiones anteriores por la constante $\frac{b}{2n}$. De lo contrario, multiplicarlas por $\frac{b}{n}$.
11. Almacenar los resultados preliminares en la posición i -ésima de dos vectores de densidades X y W , respectivamente para la expresión evaluada con α^{j-1} , y la expresión evaluada con ϕ (Estos vectores se deben crear previamente, y tienen la función de almacenar cada una de las densidades encontradas en el paso anterior).
12. Aumentar el contador tal que $i = i + 1$, y repetir los pasos desde el 8 hasta el 11, b veces.
13. Obtener las densidades en $h(\phi)$ y $h(\alpha^{j-1})$

14. Evaluar la probabilidad de aceptación, de manera tal que esta es igual a $\min \left\{ 1, \frac{h(\alpha^{j-1}) \sum_{i=1}^b W_i}{h(\phi) \sum_{i=1}^b X_i} \right\}$. Si el movimiento es aceptado, entonces $\alpha^j = \phi$, de lo contrario $\alpha^j = \alpha^{j-1}$.
15. Aumentar el contador tal que $j = j + 1$, y repetir los pasos desde el 6 al 14, tantas veces como sea el total de iteraciones que haya definido (Se recomienda un número suficientemente grande, en este caso se emplearon 10000 iteraciones) .

Así, bajo el mismo escenario de simulación con respecto al proceso de elicitación, el resultado de los números aleatorios generados para la distribución a priori de α se muestra en la Figura **6-3**, que es el histograma obtenido al llevar a cabo los algoritmos en cuestión. Allí, se evidencia que la densidad de aquellos números se asemeja en gran medida a la curva de la densidad aproximada (Figura **6-2**) para la distribución a priori de α .

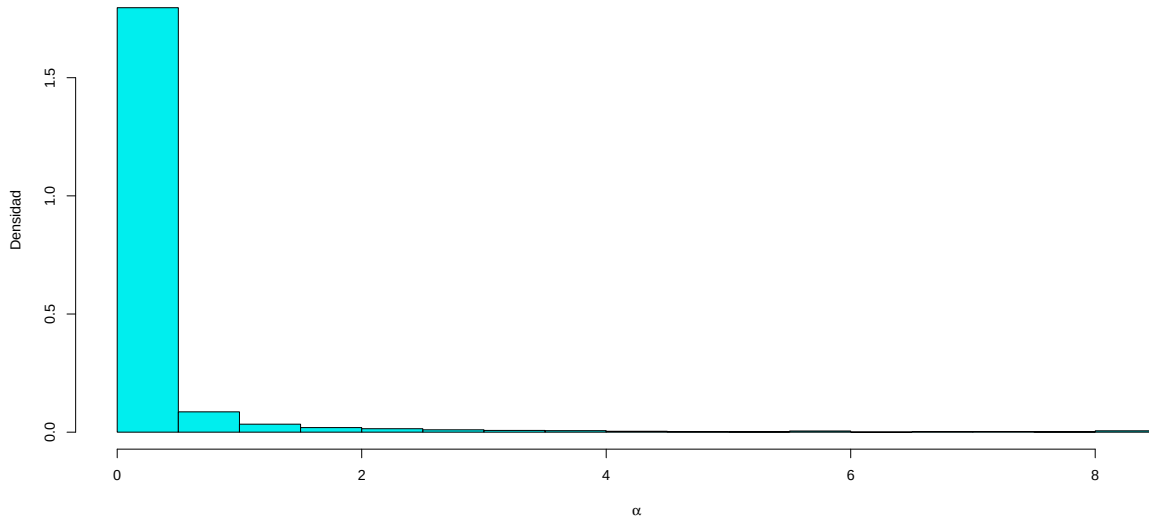


Figura 6-3: Densidad de números aleatorios obtenidos para distribución a priori de α

Continuando con el análisis, se tiene que la esperanza a priori es $E[\alpha] = \hat{\alpha} = 0,2457$, que es un valor que refleja lo acertado que puede ser esta metodología, dado que con un correcto proceso de elicitación, a priori se tiene un valor muy cercano al parámetro, que es $\alpha = 0,15$. Por último, en trabajos posteriores a este, se deberá analizar el efecto que tiene incluir la verosimilitud a la distribución a posteriori de α , y estudiar cuan precisa es la estimación.

6.2. Distribución A Priori para β

De la misma manera, después de aplicar el método de la transformada inversa, la Figura 6-4, muestra el resultado obtenido en la generación de números aleatorios para la distribución a priori de β (mostrada en la ecuación (5-51)). Donde, se generó en un escenario de simulación, una muestra de tamaño 10000, de números aleatorios a partir de una distribución uniforme, tal que $u \sim Unif(0, 1)$, y desarrollando el método mediante la ecuación (5-55), se obtuvo una muestra de valores de β comprendidos dentro del intervalo $(0, 2000)$.

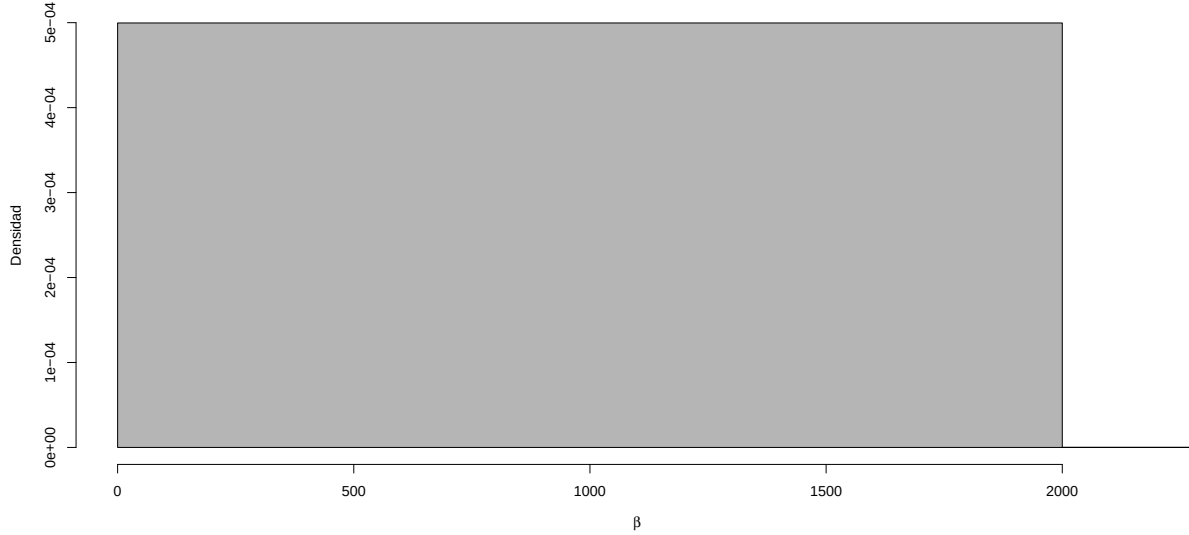


Figura 6-4: Densidad de números aleatorios obtenidos para distribución a priori de β

Esta distribución a priori no informativa, como se esperaba, muestra que en parte del espacio paramétrico de β , se asigna la misma probabilidad de ocurrencia a cualquier valor de éste. Además, si se quiere aumentar el espacio en la muestra donde tiene sentido β , solo se debe aumentar el tamaño de la muestra a la hora de generar la distribución uniforme. Además, como en el caso de la distribución a priori de α , en estudios posteriores se tendrá que analizar la precisión de las estimaciones, al incluir la verosimilitud y obtener la distribución posteriori de éste parámetro, teniendo en cuenta que esta a priori sí tiene una forma analíticamente tratable.

7 Conclusiones y Recomendaciones

Las metodología aquí planteada demuestra ser efectiva en cuanto a encontrar distribuciones a priori para los parámetros de forma de la distribución beta, pues teniendo en cuenta que éstos carecen de una interpretación práctica, plantear un proceso de elicitación para μ , que es el primer momento de dicha distribución entorno a cero, y σ^2 que es el segundo momento central, surge como una gran alternativa dado que poseen un sentido coherente en cualquier contexto de un problema dado.

También, aunque no se demuestre teóricamente, sugiere que puede existir cierta dependencia entre μ y σ^2 , pues esta última siempre es menor que la curva $\mu(1 - \mu)$, y nunca tomará un valor superior a 0.25, pues $\mu \in (0, 1)$.

Así pues, al establecer una distribución a priori conjunta para estas dos medidas, empleando conceptos del paradigma Bayesiano, es posible obtener una distribución a priori conjunta para los parámetros de forma α y β . Se obtuvo, que a partir de la última distribución mencionada, es posible hallar una distribución a priori marginal para los parámetros en cuestión.

Partiendo de ello, se puede concluir que la distribución a priori marginal del parámetro β es de forma cerrada, y tiene la característica necesaria de ser una función de densidad. Además, se le puede considerar como distribución a priori no informativa, ya que la expresión matemática de su densidad, no depende de los hiperparámetros a_1 y b_1 , referentes a la distribución a priori kumaraswamy, que se asignó a μ , por su flexibilidad para modelar variables en el espacio comprendido en el intervalo $(0, 1)$.

Del mismo modo, ya que las integrales calculadas a partir de la distribución a priori de β son analíticamente tratables, la generación de números aleatorios para ésta distribución se torna muy sencilla, ya que al hacer uso del método de la transformada inversa, se obtiene una expresión que facilita y conlleva a números coherentes al espacio paramétrico del parámetro en mención.

Por otra parte, pese que no fue posible obtener una distribución a priori marginal de forma cerrada para α , se pudo aproximar la densidad de la expresión encontrada, disponiendo de un algoritmo que emplea la regla de los trapecios, y resaltando el hecho de que es una distribución a priori informativa. Partiendo de los resultados, se concluye que la metodología expuesta en este documento, pondera de manera significativa valores para α menores que 1, teniendo en cuenta que los mismos resultados presentan densidades muy cercanas a cero, cuando se consideran cantidades de α mayores que éste valor.

Sin embargo, lo anterior no resulta en un impedimento para obtener una muestra de números aleatorios en dicha distribución, pues al combinar los pasos del algoritmo propuesto para aproximar la densidad de α y el Metropolis- Hastings, se producen números con un comportamiento similar al expuesto gráficamente por las densidades aproximadas de α . Este resultado permite, que aun sin presentar una forma analíticamente tratable, tenga sentido considerar esta distribución a priori para futuros trabajos, pues si al incluir una muestra aleatoria de la función de verosimilitud, la forma de la distribución a posteriori no cambia en gran cuantía con respecto a la priori, es posible obtener una estimación para α , empleando la misma técnica de estos algoritmos.

En este contexto, se recomienda dar continuidad en trabajos posteriores a todo el planteamiento expuesto hasta aquí, para validar si efectivamente las distribuciones a priori halladas generan resultados óptimos en la estimación de α y β . En consecuencia, una primera opción es suponer que uno de los parámetros conocidos, y así trabajar con la distribución a priori marginal correspondiente al otro parámetro y analizar los resultados para cada uno de ellos. Otra opción, es utilizar la misma estructura de la metodología, pero partiendo de distribuciones a priori para μ diferentes a la kumaraswamy, que teóricamente estén destinadas a modelar variables cuyo rango sea el intervalo $(0, 1)$. También, para investigaciones futuras se recomienda realizar un estudio de simulación donde se pueda observar el comportamiento de las estimaciones respecto a diferentes variaciones de los dos parámetros, cuyo comportamiento puede ser evaluado vía error cuadrático medio, con el fin de evaluar el método propuesto. Por último, es emplear las distribuciones marginales y condicionales a priori para estructurar una distribución posteriori conjunta, de la cual sea posible generar números aleatorios por métodos iguales o análogos a los expuestos en este trabajo.

Bibliografía

- Bautista, P. (2007), *Experimentación con un algoritmo de MCMC multiescala y autoajustable*, (Tesis Maestro en Ciencias de la Computación y Matemáticas Industriales), CIMAT, Guanajuato, Mexico.
- Box, G. & Tiao, G. (1992), *Bayesian Inference in Statistical Analysis*, Wiley, EE.UU.
- Canavos, G. (1984), *Probabilidad y estadística: aplicaciones y métodos*, Limusa, México.
- Casella, G. & Berger, R. (2002), *Statistical Inference*, Thomson Learning, EE.UU.
- Chib, S. & Ramamurthy, S. (2010), ‘Tailored randomized block mcmc methods with application to dsge models’, *Journal of Econometrics* 155(1), 19–38.
- Gamerman, D. & Lopes, H. (2006), *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Second Edition*, CRC Press, EE.UU.
- Garthwaite, P. H., Kadane, J. B. & OHagan, A. (2005), ‘Statistical methods for eliciting probability distributions’, *Journal of the American Statistical Association* 100(470), 680–701.
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A. & Rubin, D. (2014), *Bayesian Data Analysis, Third Edition*, Taylor & Francis, EE.UU.
- Johnson, N., Kotz, S. & Balakrishnan, N. (1995), *Continuous univariate distributions*, number v. 2, Wiley & Sons, New York, EE.UU.
- Kakamu, K. & Nishino, H. (2018), ‘Bayesian estimation of beta-type distribution parameters based on grouped data’, *Computational Economics* 54(1), 625–645.
- Klinger, R. (2011), *Muestreo estadístico: métodos básicos*, Universidad del Valle, Cali, Valle del Cauca, Colombia.
- Larson, R., Hostetler, R. & Edwards, B. (2006), *Cálculo con geometría analítica*, McGraw-Hill, México.
- Ma, Z. & Leijon, A. (2011), ‘Bayesian estimation of beta mixture models with variational inference’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(11), 2160–2173.
- M.C.Jones (2009), ‘Kumaraswamy distribution: A beta-type distribution with some tractability advantages’, *Statistical Methodology* 6(1), 70–81.

- Owen, C. (2008), *Parameter Estimation For The Beta Distribution*, (Thesis Master of Science), Brigham Young University, Utah, EE.UU.
- R Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- R.Tovar (2015), ‘Inferencia bayesiana e investigación en salud: un caso de aplicación en diagnóstico clínico’, *Revista Médica de Risaralda* 21(1), 9–16.
- Tanner, M. (2012), *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, Springer New York, EE.UU.
- Viertl, R. & Bernardo, J. (2003), *PROBABILITY AND STATISTICS - Volume II: Probabilistic Models and Methods Foundations of Statistics*, EOLSS Publishers Company Limited, Oxford, United Kingdom.