REVISTA VIRTUAL
Universidad Católica del Norte
CATÓLICA DEL NORTE Fundación Universitaria Pioneros en educación virtual
http://revistavirtual.ucn.edu.co
No. 48 | Mayo - Agosto, 2016
ISSN-0124-5821 (en línea)

Cómo citar el artículo
Burbano Pantoja, V.M.A.; Aldana Bermúdez, E. & Valdivieso Miranda, M.A. (2016). Conocimiento estadístico-probabilístico base para calcular integrales definidas por métodos aleatorios. *Revista Virtual Universidad Católica del Norte*, 48, 331-351.
Recuperado de http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/776/1302

Conocimiento estadístico-probabilístico base para calcular integrales definidas por métodos aleatorios*

Statistical-Probabilistic Knowledge as a Basis for Calculating Definite Integrals using Random Methods

Connaissance statistique-probabiliste comme une base pour calculer intégrales définies pour méthodes aléatoire

*Este artículo se genera de un acápite del proyecto de investigación macro titulado "Modelamiento de la tasa de fallos y la confiabilidad en términos de la DLG", incluye entre otros temas: teoría de colas y líneas de espera, simulación, conocimientos del profesor. El proyecto fue desarrollado en el Grupo de Investigación Interdisciplinario en Ciencias (GICI). Fecha de inicio: mayo de 2014, fecha de posible culminación del proyecto macro: julio de 2016. Línea de Investigación de Estadística Aplicada y Educación Estadística. Facultad de Ciencias-DIN, Universidad Pedagógica y Tecnológica de Colombia.

Indexada Publindex-Colciencias (B), Latindex, EBSCO InformationServices,Redalyc, Dialnet, DOAJ, Actualidad Iberoamericana, Índice de Revistas de Educación Superior e Investigación Educativa (IRESIE) de la Universidad Autónoma de México.

Víctor Miguel Ángel Burbano Pantoja
Licenciado en Matemáticas
Especialista en Computación para la Docencia
Especialista en Pedagogía para el Desarrollo del Aprendizaje Autónomo
Magíster en Ciencias-estadística
Doctor en Ciencias de la Educación
Docente investigador de tiempo completo en la Escuela de Matemáticas y Estadística de la Universidad Pedagógica y Tecnológica de Colombia
Coordinador del Grupo de Investigación Interdisciplinario en Ciencias GICI
victorburbanop@yahoo.es, victor.burbano@uptc.edu.co

Eliécer Aldana Bermúdez
Licenciado en Matemáticas
Especialista en Docencia Universitaria
Magíster en Administración Educativa, énfasis en Dirección
Doctor en Educación Matemática
Docente del programa de Licenciatura en Matemáticas de la Universidad del Quindío, Colombia
Coordinador del Grupo de Investigación en Educación Matemática GEMAUQ
eliecerab@uniquindio.edu.co

Margoth Adriana Valdivieso Miranda
Licenciada en Matemáticas
Magíster en Ciencias-estadística
Docente investigadora de tiempo completo en la Escuela de Matemáticas y Estadística de la Universidad Pedagógica y Tecnológica de Colombia
Integrante del Grupo GICI
mavaldiviesom@yahoo.com, margoth.valdivieso@uptc.edu.co

Recibido: 18 de julio de 2015

Evaluado: 18 de abril de 2016

Aprobado: 2 de mayo de 2016

Tipo de artículo: Investigación científica y tecnológica

Resumen

En este trabajo se presenta una forma alternativa para calcular integrales definidas por medio de métodos aleatorios. En este caso, es necesario utilizar elementos del conocimiento estadístico-probabilístico para generar números pseudoaleatorios provenientes de una distribución de probabilidad uniforme en el intervalo (0,1), los cuales como datos numéricos superarán ciertas pruebas estadísticas y serán considerados números aleatorios apropiados para realizar procesos de simulación. Luego, se usa el concepto de muestra aleatoria y la ley fuerte de los grandes números para desarrollar un algoritmo y evaluar integrales definidas de forma aproximada usando el método de Monte Carlo. Finalmente, se obtienen resultados para dos casos específicos de integrales definidas en el sentido de Riemann. Se concluye que los procedimientos desarrollados conforman una metodología y una estrategia didáctica que podrían ser utilizadas para introducir la enseñanza de procesos de simulación a nivel del bachillerato o del pregrado, la cual también puede servir para ilustrar usos concretos de la integración numérica en situaciones donde no es posible encontrar soluciones analíticas.

Palabras clave
Conocimiento estadístico-probabilístico, Estrategia didáctica, Integral definida, Métodos aleatorios.

Abstract
This article presents an alternative way for calculating definite integrals by using random methods. In this case is necessary to use elements of statistical-probabilistic knowledge in order to create pseudorandom numbers coming from an uniform probability distribution in the interval (0,1), which as numerical data will pass some statistical tests and will be considered as appropriate random numbers to perform simulation processes. The concept of random

sample and the strong law of big numbers are used in order to develop an algorithm and evaluating definite integrals in a proper way by using Monte Carlo method. Finally, results have been obtained for two specific cases of definite integrals in the sense of Riemann. We conclude that the developed processes constitute a methodology a didactical strategy which could be used as an introductory teaching on simulation processes in high school or in the university, which can also be useful to show the specific use of numerical integration in situations where is not possible to find analytic solutions.

Keywords
Statistical-Probabilistic knowledge, Didactical strategy, Definite integral, Random methods.

Résumé
Dans cet article on présente une manière pour calculer intégrales définies en utilisant méthodes aléatoires. Dans ce cas, c'est nécessaire d'utiliser éléments de la connaissance statistique-probabiliste pour générer nombres pseudo aléatoires dérivés d'une distribution de probabilité uniforme dans l'intervalle (0,1), lesquels comme données numériques surpasseront certaines tests statistiques et seront considérés comme nombres aléatoires appropriés pour réaliser processus de simulation. Après on utilise le concept de échantillon aléatoires et la loi forte des grands nombres pour développer un algorithme et évaluer intégrales définies de manière approximative en utilisant la méthode de Monte Carlo. Finalement, on a réussi des résultats pour deux cas spécifiques d'intégrales définies dans le sens de Riemann. On conclut que les procèdes développés constituent une méthodologie et une stratégie didactique qui pourraient être utilisés pour introduire l'enseignement de processus de simulation dans le lycée o dans l'université, qui peut être utile pour illustrer l'utilisation particulaire de l'intégration numérique dans situations dans lesquelles n'est possible pas de trouver solutions analytiques.

Mots-clés
Connaissance statistique-probabiliste, Stratégie didactique, Intégral définie, Méthodes aléatoires.

Introducción

En la actualidad, el conocimiento estadístico-probabilístico tiene una gran aplicación en diversos campos del conocimiento humano, entre ellos en el desarrollo de procesos de simulación. Se puede aprovechar el potencial de los computadores y el mencionado conocimiento para implementar métodos aleatorios que posibiliten encontrar soluciones aproximadas a una gran variedad de problemas que por métodos analíticos no se habían podido resolver, tal es el caso de la evaluación de cierto tipo de integrales.

En la implementación de modelos de simulación es necesario utilizar métodos aleatorios soportados en modelos matemáticos y lógicos que posibiliten de forma razonable emular y aproximar soluciones al problema de interés. Estos métodos usan elementos teóricos que forman parte del conocimiento disciplinar (Ball, Thames y Phelps, 2008) y del conocimiento didáctico del contenido (Shulman, 1986, 1987; Bolívar, 2005) a fin de lograr una mejor comprensión de la realidad y la enseñanza de determinados tópicos de interés, como las integrales definidas; en este caso, dicho conocimiento es proporcionado por la estadística matemática, la teoría de probabilidad, la programación de computadores y el diseño de algoritmos, entre otros. Después de hacer distinción entre el conocimiento especializado del contenido matemático, el conocimiento común y el horizonte matemático (Ball y Bass, 2009), se requiere proporcionar elementos didácticos que posibiliten su estructuración en el estudiante (Piaget y García, 1982;

Dubinsky, 1991, 1996; DeVries, 2001). En este trabajo se recurre al conocimiento especializado para implementar una metodología que incluye la generación de números aleatorios, el diseño de un algoritmo de cálculo y la obtención de resultados numéricos. Asimismo se hace uso del conocimiento común sobre integrales para interpretar las aproximaciones que se puedan lograr en torno a un valor que corresponda a una integral definida.

En diversos contextos concretos, la utilización de modelos de simulación es un paso necesario en la construcción del conocimiento científico y cuando se trate de la enseñanza de la probabilidad, la simulación puede constituirse en un atractivo instrumento; esta, desarrollada a través de modelos aleatorios, se constituye en una de las más destacadas ideas estocásticas para comprender la realidad (Heitele, 1975; Inzunsa, 2001; Pantoja, 2014; Burbano, Valdivieso & Salcedo, 2014; Burbano, Pinto & Valdivieso, 2015). La metodología que proponemos incluye los siguientes procedimientos: 1) Revisión de los aspectos conceptuales y procesos necesarios para evaluar integrales utilizando números aleatorios, 2) generación de una cantidad considerable de números pseudoaleatorios y aplicación de pruebas estadísticas para verificar si estos corresponden a números aleatorios y 3) diseño e implementación de un algoritmo fundamentado en la ley de los grandes números como un método para evaluar integrales definidas de manera aproximada con los números aleatorios generados o aplicación del método de Monte Carlo. Estos cálculos también podrían obtenerse por otros métodos de corte determinista como el método de los rectángulos, los trapecios o la regla de Simpson, cuando de funciones reales integrables Riemann se trate.

Elementos del marco teórico

En esta sección se desarrollan los aspectos conceptuales relacionados con el conocimiento estadístico-probabilístico necesarios para desarrollar procesos de simulación, entre ellos: métodos aleatorios, números pseudoaleatorios, distribución de probabilidad, muestra aleatoria, y la ley fuerte de los grandes números.

Métodos aleatorios

Los métodos aleatorios utilizan los conceptos de número aleatorio, variable aleatoria, distribución de probabilidad y otros que posibilitan construir modelos de simulación. El concepto de simulación se ha ido complejizando desde su formulación hecha en la segunda mitad del siglo XX (Ríos, Ríos & Jiménez, 2000). De manera formal, un sistema: S simula a T si: S y T son sistemas formales, T recoge las características de un sistema real, S se toma como una aproximación de T, el modelo S con sus reglas de validez no está exento de error (Churchman, 1973). Es posible comprender la simulación como el arte de construir modelos que permitan estudiar el comportamiento de un sistema real. Esta corresponde a un conjunto de técnicas numéricas que permite realizar experimentos en una computadora digital usando modelos matemáticos y lógicos (Naylor, Balintfy & Chu, 1977). Además, la simulación estadística ha de entenderse como un proceso consistente en colocar en correspondencia dos experimentos aleatorios diferentes, con la condición de que a cada suceso elemental del primer experimento le corresponda un suceso elemental del segundo y solo uno, de tal manera que los sucesos puestos en correspondencia en ambos experimentos tengan igual probabilidad de ocurrir (Batanero, 2001). En este contexto, los métodos aleatorios y la simulación también se podrían utilizar como recursos didácticos para enseñar conceptos de probabilidad (Burbano *et al*., 2015).

Números pseudoaleatorios

Para realizar procesos de simulación por computador, la base fundamental son los llamados números aleatorios. Estos son números que se comportan de manera no determinista como valores de variables aleatorias, de forma similar a como se pudieran comportar los fenómenos no controlables en la naturaleza (Burbano *et al*., 2014). Se han ideado diversos mecanismos y métodos para obtener secuencias de números que puedan considerarse como aleatorios. Inicialmente se generaron números aleatorios usando métodos manuales como ruletas, el método de los cuadrados medios y métodos basados en congruencias lineales, entre otros. Actualmente, se utilizan métodos computacionales para generar números aleatorios. Los números generados por computadora se denominan pseudoaleatorios y requieren superar algunas pruebas estadísticas para ser considerados como verdaderos números aleatorios.

El método de los cuadrados medios, ideado por Neumann en la década de los cuarenta del siglo XX, se considera como el primer algoritmo computacional para obtener números aleatorios. Este algoritmo permite generar pequeñas cantidades de números pseudoaleatorios ya que converge muy rápido a números cuyos dígitos son todos ceros teniéndose que elegir una nueva semilla. Actualmente, la mayoría de los algoritmos para generar números aleatorios en (0,1) se basan en el concepto de congruencia lineal. El primero en introducir las congruencias lineales en la generación de números pseudoaleatorios fue Lhemer (1951) a mediados del siglo pasado. En tal caso, dados los enteros positivos *a, b* y *n* un algoritmo para generar los números pseudoaleatorios U_i i= 1, 2, 3,..., *m* puede involucrar una congruencia lineal de la forma $X_{i+1}=(aX_i + b)$ modulo (n) que inicia con una semilla X_0. Por ejemplo, si a = 1, b = 9, se pueden generar 15 valores para X_i y 15 correspondientes a los U_i.

Sin embargo, la congruencia lineal planteada permite obtener máximo 13 números pseudoaleatorios distintos y luego empieza a repetirlos ya que los dos últimos coinciden con los dos primeros. A esta repetición que empieza en U_{14} e iría hasta U_{26} se le llama el ciclo de la congruencia lineal. Dicho ciclo continúa repitiéndose infinidad de veces, luego los primeros trece números pseudoaleatorios conforman una clase de equivalencia. Solamente si se seleccionan adecuadamente las constantes enteras y positivas *a, b* y *n,* el algoritmo alcanza un ciclo máximo de tamaño igual a *n.* Si se trabaja con computadoras de 32 bits, una buena selección de n es $n = 2^{31}-1$, $a = 7^5 = 16807$ y $b = 0$ (Saavedra & Ibarra, 2008).

En el anterior contexto, no es suficiente con buscar generadores que tengan ciclo máximo, también es deseable que los números pseudoaleatorios posean buenas propiedades estadísticas, sean fáciles de generar y de almacenar; además, que la semilla replique a la secuencia de los citados números. Un buen generador de números pseudoaleatorios en el intervalo (0,1) es el procedimiento de Kobayashi, el cual se expresa a través de $X_i = (314159269X_{i-1} + 453806245)$ módulo (2^{31}) (Law & Kelton, 1991; Glasserman, 2004). Hoy, la búsqueda de generadores de números pseudoaleatorios con mejores propiedades y más eficientes sigue abierta a la investigación (Michael, Schucany & Haas, 1976; Saavedra & Ibarra, 2008). Además, los paquetes estadísticos comerciales como SPSS, SAS y muchos otros incluyen rutinas que permiten generar automáticamente secuencias de números pseudoaleatorios a partir de determinadas semillas; por ejemplo, el software libre R.

Distribución de probabilidad y muestra aleatoria

Si X es una variable aleatoria continua con función de densidad de probabilidad *f(x)* entonces su función de distribución se denota y se define así (Lindgren, 1993; Shao, 1999; Valdivieso, 2010):

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

F(x) es una función real no decreciente continua por la derecha que cumple las siguientes tres condiciones:

i) $0 \leq F(x) \leq 1$

ii) $\lim_{x\to\infty} F(x) = 1$

iii) $\lim_{x\to-\infty} F(x) = 0$

La función de densidad de probabilidad f(x) para la variable aleatoria X satisface las dos condiciones siguientes:

i) $f(x) \geq 0$

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Ley fuerte de los grandes números

Esta ley es la más conocida de la teoría de la probabilidad. Ella establece que el promedio de una sucesión de variables aleatorias

$$X_1, X_2, \ldots, X_n \ldots$$

independientes e idénticamente distribuidas *F(x)* con media finita μ y varianza finita σ^2, converge con probabilidad 1, a la media de la distribución (Blanco, 2004). Lo anterior, se puede expresar de la siguiente manera,

$$\frac{X_1 + X_2 +, \ldots + X_n}{n} = \sum_{i=1}^{n} \frac{X_i}{n} \to \mu \quad cuando \;\; n \to \infty$$

Metodología

El trabajo se desarrolló mediante procesos de simulación asociados con el método de Monte Carlo. Estos procesos incluyen la generación de un algoritmo para su posterior ejecución en el computador, la generación de números pseudoaleatorios con sus pruebas estadísticas y el cálculo de algunas integrales definidas a partir de muestras de números aleatorios, los cuales se presentan en la sección de resultados.

Método de Monte Carlo

Las primeras aplicaciones de los métodos aleatorios por medio de los números aleatorios se hicieron en el cálculo de integrales definidas en el intervalo (0,1) (Ross, 1999). Para ello, se considera una función *f(x)* continua en (0,1) y se quiere calcular la integral *I* que se indica a continuación,

$$I = \int_0^1 f(x)dx$$

Para determinar el valor de *I*, se considera una variable aleatoria *U* con distribución uniforme en el intervalo (0,1). Luego *I* se puede expresar como el siguiente valor esperado,

$$I = E\left(f(U)\right)$$

Si $U_1, U_2, ..., U_n$ es una sucesión de variables aleatorias que conforman una muestra aleatoria proveniente de la distribución uniforme en (0,1), entonces $f(U_1), f(U_2), ..., f(U_n)$ son variables aleatorias independientes e idénticamente distribuidas con media igual a *I* (Ueberhuber, 1997). Por consiguiente, por la ley fuerte de los grandes números, se tiene con probabilidad 1 que,

$$\sum_{i=1}^{n} \frac{f(U_i)}{n} \to E\left(f(U)\right) = I \quad cuando \quad n \to \infty$$

La expresión anterior hace posible aproximar *I* al generar una gran cantidad de números aleatorios u_i y tomando como aproximación de *I* al promedio de las $f(u_i)$.

Este método para aproximar integrales, en la literatura es conocido como el método de Monte Carlo (Ueberhuber, 1997; Weinzierl, 2000). Se puede utilizar el siguiente algoritmo para calcular integrales en el sentido de Riemann, en el intervalo (0,1).

Especificar *f(x)*

Indicar *n*

Inicializar Suma=0

Para *i=1* hasta *n* hacer

Generar u_i

Suma=Suma + $f(u_i)$

Imprimir *I*=Suma/*n*

La recogida de información se realizó mediante la generación de varias secuencias de 100 números pseudoaleatorios que se sometieron a las pruebas estadísticas hasta obtener una secuencia de números aleatorios que superó esas pruebas.

Pruebas estadísticas para números pseudo aleatorios

Los números pseudoaleatorios son los generados por medio del computador o una calculadora. Para que estos puedan considerarse como números aleatorios válidos para generar valores de variables aleatorias y ser usados en procesos de simulación deben cumplir ciertas características. Si ellos pertenecen al intervalo (0,1) hipotéticamente provienen de la distribución uniforme en (0,1) y deberían ajustarse a dicha distribución, la cual presenta una media igual a ½ y una varianza igual a 1/12; además, ser estadísticamente independientes. Por lo tanto es recomendable que superen ciertas pruebas estadísticas, entre ellas: la prueba de la media, de la varianza y la de independencia (Azarang, 1996; Gentle, 1998; Kong, McCullagh & Meng, 2003; Burbano *et al*., 2014).

Para la prueba de media, se pueden contrastar las hipótesis:

$$H_0 : \mu = \frac{1}{2}$$
$$H_1 : \mu \neq \frac{1}{2}$$

O construir un intervalo de confianza para estimar la media poblacional, asumiendo que:

$$\overline{X} = \sum_{i=1}^{n} \frac{U_i}{n}$$

donde se supone que los U_i son números pseudoaleatorios provenientes de la distribución uniforme en el intervalo (0,1).

Si el (1-α)100% de confianza para estimar la media poblacional contiene a la media $\frac{1}{2} = 0.5$ entonces los números pseudoaleatorios se ajustan a una distribución uniforme en (0,1).

Para la prueba de varianza, se calcula la varianza de los números pseudoaleatorios generados, mediante la expresión:

$$S^2 = \sum_{i=1}^{n} \frac{\left(U_i - \overline{X}\right)^2}{n-1}$$

La variable aleatoria siguiente,

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

Sigue una distribución chi cuadrado con $n-1$ grados de libertad, siempre y cuando se cumpla el supuesto de normalidad para los correspondientes datos (Azarang, 1996; Kong, McCullagh & Meng, 2003; Burbano *et al*., 2014). Como se dispone del valor,

$$\sigma = \sqrt{\frac{1}{12}}$$

se puede obtener un intervalo de confianza que contenga a S^2 , a dicho intervalo se lo denota con:

(L_i, L_s), donde: $L_i = \dfrac{\chi^2_{\frac{\alpha}{2}}}{12(n-1)}$ y $L_s = \dfrac{\chi^2_{1-\frac{\alpha}{2}}}{12(n-1)}$

Una prueba de independencia permite verificar que los números pseudoaleatorios no estén correlacionados (Azarang, 1996). A continuación se describe la prueba de corridas (rachas). El proceso para aplicar la prueba de rachas arriba y abajo involucra los siguientes pasos:

Generar n números pseudoaleatorios

Clasificar cada número pseudoaleatorio teniendo en cuenta el siguiente criterio:

Si $U_i < U_{i-1}$ se coloca un signo menos (-)

Si $U_i > U_{i-1}$ se coloca un signo más (+) 340

Para $i = 2, 3, \ldots, n$. Los U_i son números pseudoaleatorios

Calcular el número *h* de corridas observadas, el número de corridas es una variable aleatoria, con:

$$E(h) = \frac{2n-1}{3} \qquad Var(h) = \frac{16n-29}{90}$$

Calcular el valor esperado y la varianza de la variable aleatoria *h*, por medio de las expresiones anteriores.

Aproximar por una variable con distribución normal estándar,

$$Z = \frac{h - E(h)}{\sqrt{Var(h)}} = \frac{h - \frac{2n-1}{3}}{\sqrt{\frac{16n-29}{90}}}$$

Si se cumple la condición $Z_{\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}}$ para un α nivel de significancia, entonces los números pseudoaleatorios son independientes.

Otra forma válida consiste en utilizar la prueba de Kolmogorov –Smirnov (K-S) para determinar si los números pseudoaleatorios se ajustan a una distribución uniforme en el intervalo (0,1); un software estadístico como SPSS presenta una opción para esta prueba (Pérez, 2005).

Resultados

Para esta investigación se generaron varias secuencias de 100 números pseudoaleatorios que se sometieron a las pruebas estadísticas. El proceso de generación se realizó hasta que se obtuvo una secuencia de 100 números que superó las pruebas estadísticas de la media, la varianza y la de independencia.

Esta secuencia se presenta en la Tabla 1., en las columnas u_i, las cuales denotan los números aleatorios obtenidos. Con estos, inicialmente se realiza el cálculo de una integral definida en el intervalo (0,1) que por métodos analíticos no se puede resolver; en este caso, utilizando los números aleatorios generados se logra una aproximación de la integral propuesta. Luego, se modificó el algoritmo planteado para el intervalo (0,1) de modo que posibilita el cálculo de integrales definidas en el intervalo (a, b). Con dicho algoritmo, se efectúa el cálculo de otra integral definida en el intervalo (-3,2) que tampoco se puede resolver por métodos analíticos.

Verificación de pruebas estadísticas para los números pseudoaleatorios

Al aplicar la prueba de Kolmogorov- Simirnov (K-S) usando un nivel de significancia del 5%, a los números pseudoaleatorios u_i de la Tabla 1., se obtuvo un P-valor de 0.443, lo cual indica que los u_i se ajustan a una distribución uniforme en el intervalo (0,1), con media ½ y varianza 1/12. Al determinar un intervalo de confianza del 95% para estimar la media de la población de donde provienen los u_i se obtiene (0.4234 , 0.5433) el cual contiene a la media poblacional $\mu = 0.5$, esto también permite aceptar la hipótesis,

$$H_0 : \mu = 0.5$$

Asimismo, mediante la prueba K-S, se verificó que los U_i cumplen el supuesto de normalidad al obtener un P-valor de 0.463. Al calcular,

$$S^2 = \sum_{i=1}^{100} \frac{\left(U_i - \overline{X}\right)^2}{n-1} = 0.091$$

Con un nivel de confianza del 95%, se determinan:

$$L_i = \frac{\chi^2_{\frac{\alpha}{2}}}{12(n-1)} = \frac{73,36}{1188} = 0,061750842 \qquad L_s = \frac{\chi^2_{1-\frac{\alpha}{2}}}{12(n-1)} = \frac{128,42}{1188} = 0,108097643$$

Luego el intervalo $(L_i, L_s) = (0.06175084\ ,\ 0.10809764)$ contiene al valor $S^2 = 0.091$ esto indica que los u_i cumplen la prueba de la varianza.

Ahora para verificar la prueba de independencia se tiene que $h = 70$,

$$E(h) = \frac{2n-1}{3} = \frac{2(100)-1}{3} = \frac{199}{3} = 66.3333$$

$$Var(h) = \frac{16n-29}{90} = \frac{16(100)-29}{90} = \frac{1571}{90} = 17.4555$$

$$Z = \frac{h-E(h)}{\sqrt{Var(h)}} = \frac{70-66.3333}{\sqrt{17.4555}} = \frac{3.6667}{4.1779} = 0.8776$$

Además, para un nivel de significancia del 5%, resulta que,

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = -1.96\ , \qquad Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$$

Luego se cumple que,

$$Z_{\frac{\alpha}{2}} < Z = 0.8776 < Z_{1-\frac{\alpha}{2}}$$

Lo anterior indica que los u_i son independientes.

Cálculo de una integral en el intervalo (0,1)

A continuación se calcula la siguiente integral,

$$I = \int_0^1 e^{-x^2} dx$$

En este caso se tiene que,

$$f(x) = e^{-x^2}$$

El algoritmo generado es el siguiente;

n=100

Suma=0

Para *i=1* hasta 100 hacer

Generar u_i

Suma=Suma + $f(u_i)$

Imprimir *I*= Suma/100=75.33/100=0.7533 aproximadamente. Ver datos de la Tabla 1.

En el algoritmo anterior, se tiene que,

$$f(u_i) = e^{-u_i^2}$$

Luego,

$$I = \int_0^1 e^{-x^2} dx \cong 0.7533$$

El cálculo de la integral anterior, definida en el intervalo (0,1) se efectuó usando el mismo algoritmo, inicialmente con los primeros 20 números aleatorios de la Tabla 1., obteniéndose un valor aproximado de 0.700. También se calculó utilizando los 50 primeros números aleatorios de la Tabla 1., y se obtuvo un valor aproximado de 0.7188 y finalmente con 100 números aleatorios se logró una aproximación de 0.7533. Se observa que a medida que se aumenta la cantidad de números aleatorios se mejora en la aproximación del cálculo de la integral a su verdadero valor.

Tabla 1. Números aleatorios y su valor en la función integrable Riemann *f(x)*

u_i	$f(u_i)$	u_i	$f(u_i)$	u_i	$f(u_i)$	u_i	$f(u_i)$
0,185	0,96635	0,03914	0,99847	0,67909	0,63055	0,89368	0,44993
0,2426	0,94284	0,81246	0,5168	0,44054	0,8236	0,65519	0,65098
0,1873	0,96553	0,55936	0,73134	0,41055	0,84489	0,00604	0,99996
0,99148	0,37417	0,60235	0,69571	0,49906	0,77953	0,06636	0,99561
0,27257	0,9284	0,7561	0,56457	0,26841	0,93049	0,03949	0,99844
0,66583	0,6419	0,10771	0,98847	0,06261	0,99609	0,00359	0,99999
0,66942	0,63883	0,37525	0,86865	0,55984	0,73094	0,21903	0,95316
0,60124	0,69664	0,74829	0,57124	0,75945	0,56171	0,19239	0,96366
0,93015	0,42098	0,9483	0,40687	0,06025	0,99638	0,95121	0,40462
0,83624	0,49693	0,44475	0,82053	0,71732	0,59777	0,6745	0,63448
0,94718	0,40773	0,71366	0,60091	0,15944	0,9749	0,55712	0,73317
0,53859	0,7482	0,57719	0,71666	0,41186	0,84398	0,57234	0,72067
0,8724	0,46716	0,04731	0,99776	0,96874	0,39123	0,09829	0,99039
0,70735	0,60632	0,52608	0,75824	0,7203	0,59522	0,21466	0,95497
0,38465	0,86247	0,35501	0,88159	0,36069	0,87801	0,92961	0,4214
0,48412	0,79107	0,47514	0,79791	0,60091	0,69691	0,82472	0,50653
0,01463	0,99979	0,97605	0,38571	0,66463	0,64292	0,82711	0,50454
0,18997	0,96455	0,67872	0,63087	0,32845	0,89774	0,61451	0,68549
0,58185	0,7128	0,70012	0,61252	0,44155	0,82286	0,33946	0,89116
0,97622	0,38558	0,02507	0,99937	0,52775	0,7569	0,93519	0,41704
0,66049	0,64646	0,01238	0,99985	0,02843	0,99919	0,23519	0,94619
0,74259	0,57612	0,78173	0,54275	0,50459	0,77522	0,82743	0,50427
0,28798	0,92041	0,88734	0,45504	0,03874	0,9985	0,26717	0,93111
0,37046	0,87176	0,53235	0,75322	0,03061	0,99906	0,09598	0,99083
0,65587	0,6504	0,19773	0,96166	0,04258	0,99819	0,13866	0,98096

Algoritmo para calcular integrales en el intervalo *(a, b)*

Para calcular integrales de la forma,

$$I = \int_a^b f(x)dx$$

Se procede a realizar la siguiente sustitución: $t = \frac{x-a}{b-a}$ de donde $(b-a)dt = dx$

Luego,

$$I = \int_a^b f(x)dx = \int_0^1 (b-a) f((b-a)t + a)dt$$

A continuación se presenta un procedimiento para determinar el valor aproximado de una integral definida en el intervalo *(a, b)* para una función integrable en el sentido de Riemann en dicho intervalo.

Cálculo de una integral en el intervalo *(a, b)*

Ahora se requiere calcular la integral dada por,

$$I = \int_{-3}^{2} 0.3989 e^{-0.5x^2} dx$$

Después de identificar la función a integrar y realizando la sustitución indicada en la sección 4.3 se tiene,

$$f(x) = 0.3989 e^{-0.5x^2}$$

$$I = \int_{-3}^{2} f(x)dx = \int_0^1 (2-(-3)) f((3-(-3))t + (-3))dt$$

$$I = \int_{-3}^{2} f(x)dx = \int_0^1 5(0.3989) f(5t-3)dt = \int_0^1 5(0.3989) e^{-0.5(5t-3)^2} dt$$

El algoritmo generado es el siguiente;

n=100

Suma=0

Para *i=1* hasta 100 hacer

Generar u_i

Suma=Suma + $5(0.3989)e^{-0.5(5u_i-3)^2}$

Imprimir *I*= Suma/100=98.18/100=0.9818 aproximadamente. Ver datos de la Tabla 2.

Luego,

$$I = \int_{-3}^{2} 0.3989 e^{-0.5x^2} dx \cong 0.9818$$

En la lectura de la Tabla 2, se debe tener en cuenta que,

$$f(u_i) = 5(0.3989) e^{-0.5(5u_i - 3)^2}$$

También en este caso, el cálculo de la integral, definida en el intervalo (-3,2) se realizó utilizando el algoritmo obtenido para el intervalo *(a,b)*; inicialmente con los primeros 20 números aleatorios de la Tabla 2., obteniéndose un valor aproximado de 0.9600. También se calculó utilizando los 50 primeros números aleatorios de la Tabla 2., y se obtuvo un valor aproximado de 1.07 y finalmente con 100 números aleatorios se logró una aproximación de 0.9818. Se infiere que a medida que se aumenta la cantidad de números aleatorios se mejora en la aproximación del cálculo de la integral hacia su verdadero valor.

Tabla 2. Números aleatorios y su valor en la función integrable Riemann *f(x)*

u_i	$f(u_i)$	u_i	$f(u_i)$	u_i	$f(u_i)$	u_i	$f(u_i)$
0,185	0,23168	0,03914	0,0391	0,67909	1,84449	0,89368	0,67861
0,2426	0,40402	0,81246	1,13446	0,44054	1,45143	0,65519	1,91999
0,1873	0,23725	0,55936	1,95375	0,41055	1,27348	0,00604	0,02425
0,99148	0,29367	0,60235	1,99436	0,49906	1,75599	0,06636	0,05674
0,27257	0,52218	0,7561	1,4708	0,26841	0,50459	0,03949	0,03929
0,66583	1,88933	0,10771	0,09643	0,06261	0,05396	0,00359	0,02338
0,66942	1,8779	0,37525	1,06076	0,55984	1,95469	0,21903	0,32503
0,60124	1,99446	0,74829	1,51516	0,75945	1,45149	0,19239	0,24997
0,93015	0,51064	0,9483	0,43779	0,06025	0,05228	0,95121	0,42679
0,83624	0,9928	0,44475	1,47567	0,71732	1,67924	0,6745	1,86082
0,94718	0,44207	0,71366	1,69708	0,15944	0,17626	0,55712	1,94918
0,53859	1,90266	0,57719	1,98157	0,41186	1,28138	0,57234	1,97552
0,8724	0,78889	0,04731	0,04381	0,96874	0,3645	0,09829	0,08578
0,70735	1,72693	0,52608	1,86282	0,7203	1,66445	0,21466	0,31171
0,38465	1,11706	0,35501	0,9419	0,36069	0,97485	0,92961	0,51292
0,48412	1,68631	0,47514	1,64135	0,60091	1,99448	0,82472	1,06094
0,01463	0,02752	0,97605	0,34052	0,66463	1,89303	0,82711	1,04671

0,18997	0,24386	0,67872	1,84584	0,32845	0,79346	0,61451	1,98926
0,58185	1,9863	0,70012	1,75961	0,44155	1,45727	0,33946	0,85375
0,97622	0,33998	0,02507	0,03202	0,52775	1,86851	0,93519	0,48968
0,66049	1,90533	0,01238	0,02663	0,02843	0,0336	0,23519	0,37788
0,74259	1,54689	0,78173	1,31992	0,50459	1,77999	0,82743	1,04481
0,28798	0,59063	0,88734	0,71059	0,03874	0,03888	0,26717	0,49942
0,37046	1,0323	0,53235	1,8836	0,03061	0,03466	0,09598	0,08332
0,65587	1,91818	0,19773	0,26385	0,04258	0,04103	0,13866	0,13945

Discusión de resultados

La generación de números aleatorios es un proceso dispendioso que se debe realizar cuidadosamente. Para ello, es necesario realizar procedimientos de generación de secuencias de números pseudoaleatorios y someterlas a las pruebas estadísticas hasta que alguna de ellas supere las pruebas de la media, la varianza y la de independencia, a fin de garantizar estadísticamente que la secuencia corresponda a una secuencia de números aleatorios. Hoy, con la presencia del computador y de software especializado, esta tarea puede resultar muy fácil y rápida. Pero las personas deben ser conscientes de que dichos procedimientos han de involucrar el conocimiento estadístico-probabilístico y los demás que se requieren en el desarrollo de procesos de simulación.

En las tablas 1 y 2, se presentan los números aleatorios u_i y sus respectivos valores $f(u_i)$ correspondientes a la función que en cada caso se pretendía integrar. Para su obtención se elaboró un segmento de programa para computador en un lenguaje de alto nivel.

Los resultados obtenidos sobre el cálculo de una integral definida de manera aproximada en el intervalo (0,1) y otra en el intervalo (-3,2) muestran que por medio de los métodos aleatorios se pueden obtener buenas aproximaciones para integrales que por métodos analíticos no se pueden resolver. Se puede comprobar que si se aumenta la cantidad de números aleatorios generados, se consiguen mejores aproximaciones hacia el verdadero valor de la integral propuesta; esto se explica por la convergencia de variables aleatorias en que se soporta la ley de los grandes números y el método de Monte Carlo. También de manera empírica, se observó que a medida que se pasa de 20 a 50 y luego a 100 números aleatorios, se obtuvieron mejores aproximaciones en los dos casos presentados.

Los procedimientos desarrollados en el presente trabajo se pueden utilizar para calcular, de modo aproximado, una gama amplia de integrales definidas en cualquier intervalo *(a, b)*, incluyendo también aquellas que sí se pueden evaluar por los métodos tradicionales. Para ello, se recomienda realizar procesos y aplicar algoritmos similares a los aquí presentados. La secuencia de pasos desarrollados puede constituirse en una herramienta metodológica y una guía didáctica para aproximar integrales definidas por métodos aleatorios, cuyo fundamento se encuentra en el método de Monte Carlo.

Los resultados obtenidos no se hubieran logrado sin el uso del computador con un buen generador de números aleatorios y sin el soporte teórico dado por el conocimiento estadístico-probabilístico relacionado con la generación de números aleatorios, la convergencia de variables aleatorias, la ley de los grandes números y el método de Monte Carlo; este último también tuvo su fundamento en el mencionado conocimiento. Cuando los procesos de simulación se realizan de manera automatizada en el computador mediante software especializado, por lo general, las personas no alcanzan a dimensionar los tipos de conocimientos que en realidad se requieren para la implementación de dichos procesos, teniéndose la creencia de que el computador los realiza sin necesidad de que la gente tenga ningún conocimiento teórico pertinente.

Conclusiones

Una metodología y estrategia didáctica para calcular integrales definidas a través de métodos aleatorios incluye elementos como el concepto de integral, generación de números aleatorios con superación de ciertas pruebas estadísticas, implementación del algoritmo basado en la ley de los grandes números o aplicación del método de Monte Carlo.

Para el desarrollo de procesos de simulación basados en métodos aleatorios, como los indicados en el presente trabajo, es necesario integrar el conocimiento estadístico-probabilístico con algunas herramientas de tipo computacional y el diseño e implementación de algoritmos apropiados, algunos de ellos con fundamento en el método de Monte Carlo.

Los resultados obtenidos sobre el cálculo de una integral definida en el intervalo (0,1) y otra en el intervalo *(-3,2)* muestran las bondades y las potencialidades de los métodos aleatorios sobre el cálculo de integrales que por métodos analíticos no se pueden resolver y la extensión de los procedimientos respectivos a integrales de funciones continuas definidas en cualquier intervalo *(a, b)*, las que han de ser integrables Riemann en el mencionado intervalo.

Se pueden obtener mejores aproximaciones en el cálculo de integrales definidas por métodos aleatorios si en los algoritmos implementados se utilizan mayores cantidades de números aleatorios que hayan superado las pruebas estadísticas.

Cada vez que los procesos de simulación se ejecutan de forma automática en el computador mediante software especializado, por lo general, los individuos difícilmente dimensionan los tipos de conocimientos requeridos para la implementación de estos procesos, teniéndose la creencia de que el computador los realiza sin necesidad de que las personas tengan ningún conocimiento teórico pertinente.

Las personas han de ser conscientes de que los computadores solo realizan los cálculos referentes a los procesos de simulación, de la forma como se les indique a través de órdenes entendibles a la máquina y que los seres humanos usando una base de conocimientos adecuada son quienes diseñan y programan dichos procesos en concordancia con el interés que se tenga en modelar un determinado fenómeno aleatorio o resolver un particular problema; tal es el caso de obtener aproximaciones al valor correspondiente a una integral definida en el sentido de Riemann que por métodos analíticos no se haya podido resolver.

Referencias

Azarang, M. R. (1996). *Simulación y análisis de modelos estocásticos.* México: McGraw -Hill.

Ball, D. L.; Thames, M. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.

Ball, D. L. & Bass, H. (2009). With an eye on the mathematical horizon: Knowing mathematics for teaching to learners' mathematical futures. Trabajo presentado en *43rd Jahrestagung der Gesellschaft für Didaktik der Mathematik*. Oldenburg, Alemania. Disponible en www.mathematik.tu-dortmund.de/ieem.

Batanero, C. (2001). *Didáctica de la estadística*. Granada: Grupo de Investigación en Educación Estadística. Universidad de Granada.

Blanco, L. (2004). Probabilidad. Universidad Nacional de Colombia.

Bolívar, A. (2005). Conocimiento didáctico del contenido y didácticas específicas. *Profesorado*, *9*, 2.

Burbano, V.M.A. et al. (2014). *Simulación con modelos aleatorios: Conocimiento estadístico-probabilístico y simulación*. Tunja: Editorial Uptc.

Burbano, V. M. Á., Pinto, J. E. & Valdivieso, M. A. (2015). Formas de usar la simulación como un recurso didáctico. *Revista Virtual Universidad Católica del Norte*, *2*(45), 16-37

Churchman, C. W. (1973). *El enfoque de sistemas*. México: Diana.

DeVries, D. J. (2001). RUMEC / APOS Theory *Glossary. Georgia Collage & State University.Milledgeville. http://www.cs.gsu.edu/~rumec/Papers/glossary.html*.

Dubinsky, E. (1991). Reflective Abstraction in Advanced Mathematical Thinking, en Tall, D. (ed.). *Advanced Mathematical Thinking*. Dordrecht, The Netherlands: Kluwer Academic Press, pp. 95-123.

Dubinsky, E. (1996). Aplicación de la perspectiva piagetiana a la educación matemática universitaria. *Educación Matemática*, 8(3), pp. 24-41.

Gentle, J. E. (1998). *Random number generation and Monte Carlo methods.* New York: Springer.

Glasserman, P. (2004). *Monte- Carlo Methods in Financial Engineering.* Computational Finance. New York: Springer Verlag.

Heitele, D. (1975). An epistemogogical view and fundamental stochastic ideas. *Educational studies in Mathematics*, 6, 187 -205.

Inzunsa, S. (2001). *Una propuesta didáctica para la enseñanza de la probabilidad en el bachillerato basada en el enfoque frecuencial y simulación computacional* (Doctoral dissertation, Tesis de Maestría inédita, Departamento de Matemática Educativa, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México).

Kong, A. McCullagh, P. & Meng, X. L.(2005). A theory of statistical models for Monte Carlo Integration. *J. R. Statist. Soc. B 65*, Part 3, pp. 585–618

Law, M. A. M. & Kelton, W. D. (1991). *Simulation Modeling and Analysis.* México: McGraw Hill.

Lehmer, D. H. (1951). Mathematical methods in large-sacle computing units. *Ann. Comput Lab. Harvard Univ, 26*, 141 - 146

Lindgren, B. (1993): *Statistical Theory*, Fourth Edition, Editorial Chapman Hall. United of States of América.

Mayorga, H. (2003). *Inferencia estadística*. Notas de clase, Unibiblos. Universidad Nacional de Colombia.

Michael, J. R., Schucany, W. R. & Haas, R. W. (1976). Generating random variates Using transformation whit multiple roots. *The American Statistician, 30*, 88 -90.

Naylor, Th., Balintfy, J. & Chu, K. (1977). *Técnicas de simulación en computadoras.* México: Ed. Limusa.

Pantoja, V. M. Á. B. (2014). Simulación en el contexto de la didáctica de la estadística. *ALAMMI, Revista científica*, (2).

Papoulis, A. (1991). *Probability, Random variables and Stochastic Process*. New York: McGraw-Hill Inc.

Pérez, C. (2005). *Técnicas estadísticas con SPSS 12 Aplicaciones al análisis de datos*. Madrid: Pearson-Prentice hall.

Piaget, J. & García, R. (1982). *Psicogénesis e Historia de la Ciencia*. México, España, Argentina, Colombia. (Madrid): Siglo XXI.

Ríos, D., Ríos, S. & Jiménez, J. (2000). *Simulación, métodos y Aplicaciones.* Santafé de Bogotá: Ed. Alfa-omega.

Ross, Sh. (1999). *Simulación*. Prentice Hall. United States of America.

Saavedra, P. & Ibarra, V. (2008). *Simulación numérica de variables aleatorias*. Univ. Autónoma Metropolitana-Iztapalapa.

Shao, J. (1999). *Mathematical Statistics*. New York: Springer.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Research*, *15* (2), 4-14.

Shulman, S. (1987). Knowledge and teaching: foundations of the new reforms. *Harvard Educational Review, 57*(1), 1-22.

Ueberhuber, C. W. (1997). “Monte Carlo Techniques” in Numerical Computation 2: Methods, Software and Analysis. Berlin: Springer-Verlag, pp. 124-125 and 132-138.

Valdivieso, M. A. (2010). *Probabilidad básica y distribuciones. Apoyo al estudio independiente*. Tunja: Impresiones Jotamar.

Weinzierl, S. (2000). “Introduction to Monte Carlo Methods”. Recuperado de http://arxis.org//abs/hep-ph/0006269