

Eficiencia del sistema de propulsión WARP de la nave espacial USS Enterprise

Lucila Stancato, *I.T.B.A*, Damián Modernell, *I.T.B.A*, Juan Brasca, *I.T.B.A*, Conrado Negro, *I.T.B.A*

Resumen—Analizamos la eficacia en la generación de números pseudoaleatorios del generador de L'Ecuyer aplicando los tests χ^2 y KS , en los que determinamos la conveniencia de usar dicho generador. Utilizando realizaciones de distintas variables pseudoaleatorias, estimamos el tiempo medio de vuelo de la nave espacial USS Enterprise.

Palabras clave—Generador de L'Ecuyer, números pseudoaleatorios, propulsor WARP

1 INTRODUCCIÓN

Los dispositivos de cómputo no pueden generar números aleatorios dado que son dispositivos deterministas. La única manera de obtenerlos sería a través de un dispositivo que detecte procesos naturales como por ejemplo el intervalo de tiempo entre dos partículas α en una muestra radiactiva.

En la sección 2 utilizamos y sometemos a prueba el generador propuesto por L'Ecuyer para la generación de números pseudoaleatorios que a diferencia de los números aleatorios, pueden ser generados por una computadora. Usamos las pruebas de χ^2 y de Kolmogorov-Smirnov para determinar la bondad de ajuste del generador. En la sección 3 utilizamos una variable pseudoaleatoria generada con el algoritmo de L'Ecuyer para obtener variables pseudoaleatorias con distintas funciones de distribución. En la sección 4 calculamos la media y la varianza muestral del tiempo de vuelo del USS Enterprise.

Comprobamos la efectividad del generador de L'Ecuyer en la generación de números pseudo-aleatorios, ya que los números obtenidos con el mismo, aparentan ser aleatorios. Con dichos números generados podemos realizar simulaciones de variables más complejas como la que vemos en la sección 4.

2 GENERADOR DE L'ECUYER

El generador propuesto por L'Ecuyer en 1998 combina dos generadores lineales congruenciales (LCGs). El algoritmo está compuesto por cinco pasos:

- PASO 1: Seleccionar una semilla $X_{1,0}$ en el rango $[1, 2147483562]$ para el primer LCG y otra $X_{2,0}$ en el rango $[1, 2147483398]$ para el segundo LCG.
- PASO 2: Evaluar cada generador individual:

$$X_{1,n+1} = 40014X_{1,n} \bmod 2147483563$$

$$X_{2,n+1} = 40692X_{2,n} \bmod 2147483399$$

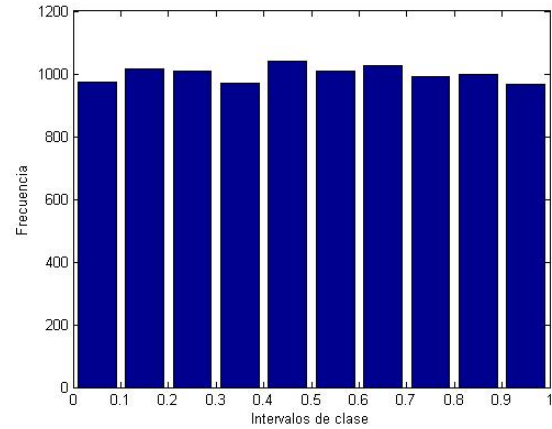


Figura 1. 10000 números generados con el generador de L'Ecuyer divididos en 10 intervalos de clase que muestran una distribución uniforme

- PASO 3: Computar:

$$X_{n+1} = (X_{1,n+1} - X_{2,n+1}) \bmod 2147483562$$

- PASO 4: Computar:

$$U_{n+1} = \begin{cases} \frac{X_{n+1}}{2147483563} & X_{n+1} > 0 \\ \frac{2147483562 - X_{n+1}}{2147483563} & X_{n+1} = 0 \end{cases}$$

- PASO 5: Hacer $n = n + 1$ y volver al PASO 2.

Los números obtenidos por medio de este generador tienen una distribución uniforme. Generamos 10000 números pseudoaleatorios con el algoritmo de L'Ecuyer. Tomamos como semillas $X_{1,0} = 23$ y $X_{2,0} = 23$, sabiendo que lo importante es no inicializar al generador en cero. Una vez obtenidas 10000 realizaciones, las dividimos en 10 intervalos de clase de amplitud uniforme, y presentamos las frecuencias obtenidas en el histograma de la figura 1.

Analizamos gráficamente las realizaciones obtenidas con el algoritmo de L'Ecuyer para determinar la dependencia de una realización con la anterior. En la figura

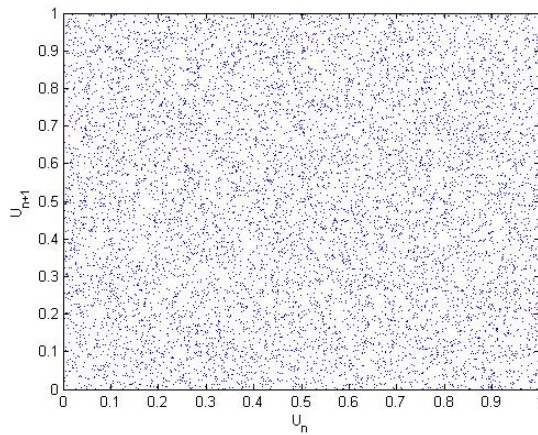


Figura 2. Observamos la distribución uniforme de los 10000 números obtenidos con el generador de L'Ecuyer en el plano

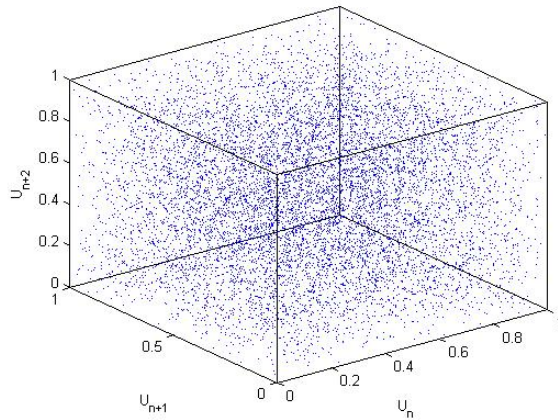


Figura 3. Observamos la distribución uniforme de los 10000 números obtenidos con el generador de L'Ecuyer en el espacio

2 mostramos en el plano las duplas (U_i, U_{i+1}) , que no dejan ver a simple vista los planos de dependencia entre una iteración y la siguiente. En la figura 3 mostramos en el espacio las ternas (U_i, U_{i+1}, U_{i+2}) , que tampoco evidencian áreas vacías, o áreas con mayor concentración de puntos que otras.

2.1 Prueba χ^2

En las figuras 2 y 3 observamos que la distribución parece ser uniforme, pero podemos usar tests más elaborados para corroborar esos resultados. El test χ^2 determina, con un nivel de significación α (que fijamos en 5%) si es razonable suponer que la distribución observada de las 10000 muestras es consistente con que la variable tenga una distribución uniforme. Las hipótesis del test son:

- $H_0: \chi_0^2 < \chi_{n-1, \alpha}$ (Los números obtenidos mediante el método de L'Ecuyer están uniformemente distribuidos)
- $H_1: \chi_0^2 \geq \chi_{n-1, \alpha}$ (Los números obtenidos mediante el método de L'Ecuyer no están uniformemente distribuidos)

Para la prueba decidimos tomar 10 intervalos de clase, determinando así 9 grados de libertad. Para estos parámetros, obtenemos de tablas el valor crítico $\chi_{n-1, \alpha}^2 = 16.919$. El estadístico χ_0^2 se computa mediante la fórmula de la ecuación 1

$$\chi_0^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

y resulta $\chi_0^2 = 5.592$ por lo que se acepta la hipótesis nula H_0 .

2.2 Prueba de Kolmogorov-Smirnov

Las hipótesis utilizadas para esta prueba son:

- $H_0: D < D_\alpha$ (Los números obtenidos mediante el método de L'Ecuyer están uniformemente distribuidos)
- $H_1: D \geq D_\alpha$ (Los números obtenidos mediante el método de L'Ecuyer no están uniformemente distribuidos)

Donde D es el resultado de la prueba y D_α es el valor crítico correspondiente a los parámetros de la prueba. Computamos el valor D con la fórmula de la ecuación 2.

$$D = \max(D^+, D^-) \quad (2)$$

siendo

$$D^+ = \max\left(\frac{i}{n} - x_i\right) \quad (3)$$

$$D^- = \max\left(x_i - \frac{i-1}{n}\right) \quad (4)$$

donde x_i es el i -ésimo valor de los calculados por el método de L'Ecuyer; y el valor i es la cantidad de realizaciones menores que x_i de la muestra de 10000. Realizando los cálculos, resulta $D = 0.0034$ y como $D_\alpha = 0.0136$, se acepta la hipótesis nula.

3 GENERACIÓN DE VARIABLES ALEATORIAS CON DISTINTAS FUNCIONES DE PROBABILIDAD

Teniendo una realización de una variable aleatoria con distribución $U[0, 1]$, podemos generar otra variable que tenga una función de distribución $F(x)$. Para hacerlo, calculamos la imagen inversa $x_i = F^{-1}(u_i)$.

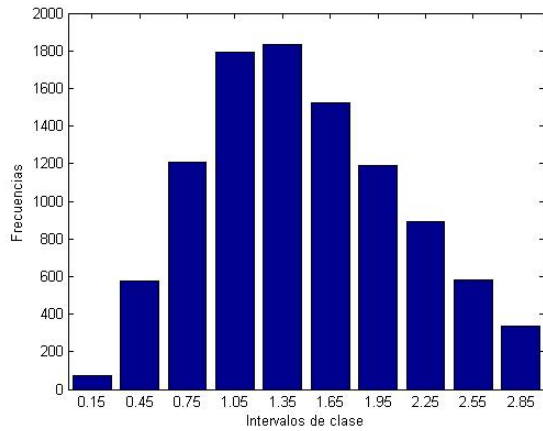


Figura 4. Histograma de una realización de una variable pseudoaleatoria con distribución de probabilidad triangular con parámetros $a=0$, $b=1$ y $c=3$; obtenida a partir de una realización de una variable con distribución uniforme.

3.1 Distribución triangular

La función de la ecuación 5 corresponde a la función densidad de probabilidad de una variable aleatoria con distribución triangular.

$$f_X(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & a \leq x \leq b \\ \frac{2(c-x)}{(c-b)(c-a)} & b < x \leq c \\ 0 & \text{Otro caso} \end{cases} \quad (5)$$

Integramos y calculamos la función inversa para obtener la expresión de la ecuación 6

$$x_i = \begin{cases} a + \sqrt{u_i(b-a)(c-a)} & 0 \leq u_i \leq \frac{b-a}{c-a} \\ c - \sqrt{(1-u_i)(c-a)(c-b)} & \frac{b-a}{c-a} < u_i \leq 1 \\ 0 & \text{Otro caso} \end{cases} \quad (6)$$

Aplicamos esta función de transformación al set de 10000 valores obtenidos en la sección 2, y obtenemos 10000 realizaciones de una variable pseudoaleatoria con distribución triangular. Dividimos las realizaciones en 10 intervalos de clase, y graficamos las frecuencias obteniendo el histograma que mostramos en la figura 4.

3.2 Distribución exponencial

La función de la ecuación 7 corresponde a la función densidad de probabilidad de una variable aleatoria con distribución exponencial.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & 0 \leq x \\ 0 & \text{Otro caso} \end{cases} \quad (7)$$

Integramos y calculamos la función inversa para obtener la expresión de la ecuación 8

$$x_i = \begin{cases} 0 & u_i = 0 \\ -\ln(u_i) & \text{Otro caso} \end{cases} \quad (8)$$

3.3 Distribución uniforme parametrizada

Para convertir una variable aleatoria con distribución uniforme $U[0,1]$ en otra variable aleatoria con distribución $U[a,b]$, usamos la expresión de la ecuación 9.

$$x_i = (b-a)u_i + a \quad (9)$$

4 ESTIMACIÓN DEL TIEMPO DE VUELO MEDIO DEL USS ENTERPRISE

La nave espacial USS Enterprise, famosa por las películas de Star Trek, es impulsada por un sistema de propulsión WARP, que se conforma por dos motores a los laterales, uno a babor y otro a estribor. Ambos propulsores son alimentados por un núcleo WARP donde se llevan a cabo reacciones de aniquilación materia / antimateria, moderadas por cristales de dilitio. Dichos cristales son procesados en una cámara controlada, llamada Matriz de Dilitio.

El tiempo de operación de cada propulsor es una variable aleatoria de distribución exponencial con tiempo medio de 10 días. A su vez, el tiempo de operación entre fallos del núcleo WARP es de 72 horas, pudiendo variar linealmente hasta en 12 horas. La nave puede propulsarse a velocidad WARP con un solo motor funcionando. Por último, las cámaras de dilitio funcionan de forma tal que, la cámara principal tiene un tiempo de operación entre 20 y 50 horas, de distribución uniforme. La cámara redundante tiene un tiempo entre fallos de 5 a 12 horas.

Definimos dos variables aleatorias $X_1 \sim U[20,50]$ horas y $X_2 \sim U[5,12]$, las cuales tienen una distribución de probabilidad uniforme y corresponden a la cámara principal, y a la cámara redundante. También definimos una variable aleatoria X_3 de distribución triangular, que corresponde al núcleo WARP, con parámetros $a = 60$ horas, $b = 72$ horas y $c = 84$ horas. Además definimos las variables aleatorias X_4 y X_5 de distribución exponencial con $\lambda = 1/240$, que corresponden a los propulsores WARP del Enterprise.

Para generar las variables pseudoaleatorias X_1, X_2, X_3, X_4 y X_5 utilizamos 5 variables pseudoaleatorias generadas a partir del algoritmo de L'Ecuyer. Para esas 5 variables tomamos 5 semillas distintas (distintas a cero), cuyos valores mostramos en la tabla 1.

Para estimar el tiempo de vuelo del USS Enterprise, definimos una variable aleatoria T que se caracteriza por la ecuación 10.

$$T = \min\{\max\{X_1, X_2\}, X_3, \max\{X_4, X_5\}\} \quad (10)$$

Tabla 1
Semillas de las variables pseudoaleatorias

Variable	Semilla
u_1	23
u_2	2
u_3	5
u_4	17
u_5	7

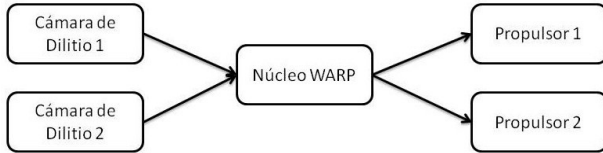


Figura 5. Esquema de representación del sistema de propulsión de la nave espacial USS Enterprise. La cámara de Dilutio 1 es la cámara principal, representada por X_1 . La cámara de Dilutio 2 es la redundancia, representada por X_2 . El núcleo WARP está representado por X_3 y los propulsores están representados por X_4 y X_5 .

Estimamos el tiempo medio de vuelo con un error porcentual menor al 5%, donde el error porcentual se computa como $100L/\bar{T}$, siendo L la longitud del intervalo de confianza. Hacemos 279 simulaciones y obtenemos una media muestral del tiempo de vuelo de $\bar{T} = 34.1932$ horas y un desvío muestral de 8.9 horas. En la figura 6 mostramos la media muestral en función de la cantidad de iteraciones realizadas. Podemos ver que se estabiliza mientras más simulaciones se hacen. El desvío muestral, que mostramos en la figura 7 disminuye a medida que aumenta la cantidad de simulaciones. A su vez, podemos ver en la figura 8, que el error porcentual, también disminuye con la cantidad de iteraciones, y que alcanza un nivel de 5% a las 279 simulaciones.

5 CONCLUSIÓN

Determinamos que para que un generador de números pseudoaleatorios sea eficaz, debe superar no uno, sino varios tests a los cuales es sometido. Quizas un generador pase un test, pero luego al obtener un gráfico en dos o tres dimensiones de una cantidad de puntos determinada, podamos ver la mala distribución de los mismos. Como vemos, el generador de L'Ecuyer pasa todos los tests. Sin embargo el haber superado las pruebas, no significa que los números que genera sean aleatorios. Al ser obtenidos con un dispositivo determinista, dichos números son generados mediante un algoritmo, lo cual contrasta con el hecho de que los números aleatorios son generados por hechos naturales.

BIBLIOGRAFÍA

- [1] Díaz, A. R. *El concepto de control*. Departamento de Ingeniería Informática. Instituto Tecnológico de Buenos Aires.

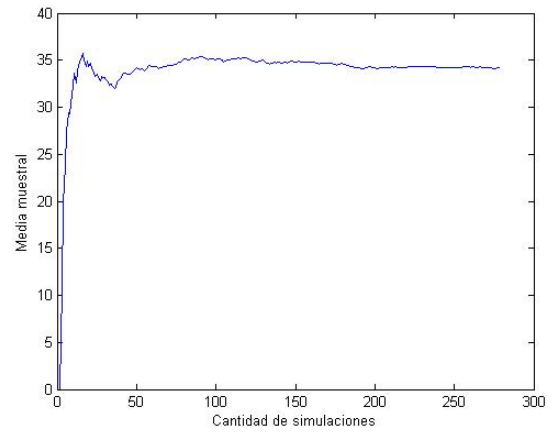


Figura 6. Media muestral del tiempo de vuelo en función de la cantidad de simulaciones.

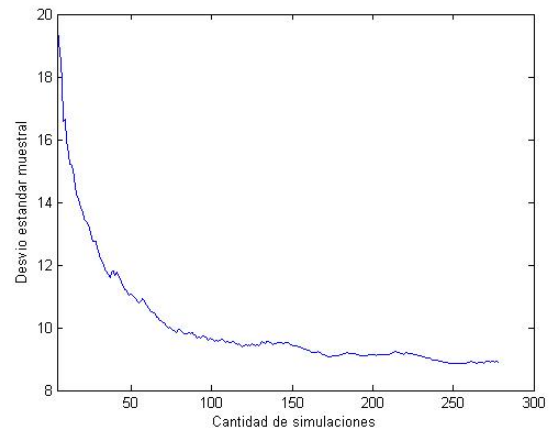


Figura 7. Desvío muestral del tiempo de vuelo en función de la cantidad de simulaciones.

- [2] Tabla de valores críticos de la distribución χ^2 . <http://www.mat.uda.cl/hsalinas/cursos/2008/probabilidad/TablaChiCuadrada.pdf>
- [3] Kolmogorov-Smirnov Test. <http://www.eridlc.com/onlinetextbook/index.cfm?fuseaction=textbook.appendix&FileName=Table7>
- [4] Tabla de distribución T-Student. http://www.dm.uba.ar/materias/probabilidades_estadistica_C/2008/2/tdist.pdf

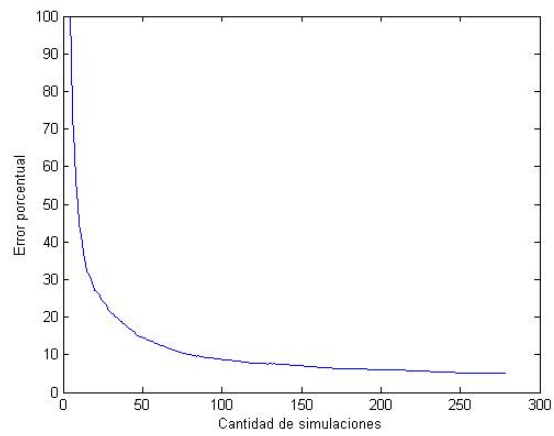


Figura 8. Error porcentual del tiempo de vuelo en función de la cantidad de simulaciones.