Simulación ITAM
31 de enero de 2018 Jorge de la Vega
2018-I

Tarea 1.

La fecha de entrega es el **16 de febrero de 2018**.

Lecturas

- Robert & Casella Capítulo 2 sección 2.1
- Dagpunar Capítulo 2
- Good random number generators are (not so) easy to find
- Linear Congruential Generator in R

Problemas

1. Probar por inducción matemática que para un generador lineal congruencial (GLC),

$$Z_i \equiv \left[a^i Z_0 + c\frac{a^i - 1}{a - 1}\right] \mod m$$

2. Calcular el periodo de $Z_i \equiv (5Z_{i-1} + 3) \mod 31$.

3. Sin calcular explícitamente ninguna Z_i, determinar cuál de los siguientes GLC's mixtos tienen periodo completo.

   (a) $Z_i \equiv [13Z_{i-1} + 13] \mod 16$.

   (b) $Z_i \equiv [12Z_{i-1} + 13] \mod 16$.

   (c) $Z_i \equiv [25437Z_{i-1} + 35421] \mod 2^{10}$.

   (d) $Z_i \equiv [Z_{i-1} + 12] \mod 13$.

   (e) El glc con parámetros: $a = 2{,}814{,}749{,}767{,}109$, $c = 59{,}482{,}661{,}568{,}307$, $m = 2^{48}$.

4. Mostrar que el promedio de las U_i's tomadas de un ciclo completo de un GLC de periodo completo es $\frac{1}{2} - \frac{1}{2m}$.

5. Dada una sucesión $X_1, X_2, \ldots, X_n$ de $\mathcal{U}(0,1)$ números pseudoaleatorios, podemos hacer una gráfica de dispersión de puntos de (X_i, X_{i+1}) para $i = 1, \ldots, n-1$ para verificar si hay independencia. Hacer esta gráfica para el GLC con parámetros $m = 1{,}024, a = 401, c = 101$ y para el GLC $m = 2^{32}, a = 1{,}664{,}525, c = 1{,}013{,}904{,}223$.

6. Probar que la parte fraccional de la suma de uniformes [0,1] $U_1 + U_2 + \cdots + U_k$ es también uniforme en el intervalo [0,1]

7. Un generador de Fibonacci obtiene X_{n+1} a partir de X_n y X_{n-1} de la siguiente forma:
$$X_{i+1} \equiv (X_i + X_{i-1}) \mod m$$
donde X_0 y X_1 están especificados.

   Supongan que $m = 5$. Sólo dos ciclos son posibles. Encontrarlos, así como su respectivo periodo.

8. Genera $10{,}000$ números con una semilla de $Z_0 = 1$ usando el generador $Z_n = 7^5 Z_{n-1} \mod (2^{31} - 1)$ Clasifica los números en 10 celdas de igual tamaño y prueben por uniformidad usando la prueba χ^2 con un nivel de confianza del 90 %. Aplicar también la prueba de rachas.

9. El método del cuadrado medio de John von Neumann es el siguiente: comenzando con $Z_0 \in \{0, 1, \ldots, 99\}$, definir Z_n para $n \in \mathbb{N}$ a ser los dos dígitos de enmedio del número de 4 dígitos Z_{n-1}^2. Si Z_{n-1}^2 no tiene 4 dígitos, se le pegan a la izquierda con ceros. Por ejemplo, si $Z_0 = 64$, tenemos que $Z_0^2 = 4096$ y entonces $Z_1 = 09 = 9$. En el siguiente paso, encontramos que $Z_1^2 = 81 = 0081$, así que $Z_2 = 08 = 8$.

   - Escriban una función que calcule Z_n a partir de Z_{n-1}.
   - La salida del cuadrado medio tiene bucles. Por ejemplo, una vez que $Z_N = 0$, tendremos que $Z_n = 0$ para toda $n \geq N$. Escriban un programa que encuentre todos los ciclos del método del cuadrado medio y lístenlos.
   - Comenten sobre la calidad del método como generador de números aleatorios.

10. Generar 15 números usando la semilla $Z_0 = 1$ del siguiente generador: $Z_n = (5Z_{n-1} + 1) \mod 16$. Hacer una prueba de Kolmogorov-Smirnov al 95 % de confianza.

11. La página PI DAY (http://www.piday.org/million/) contiene el primer millón de dígitos de π. Considerando estos dígitos:

- Realizar un histograma y verificar la hipótesis de que los dígitos corresponden a una distribución uniforme discreta.
- Verificar independencia de los dígitos, considerando las pruebas de gaps, de poker y de rachas.

  Una idea de ver los datos está en la siguiente imagen:

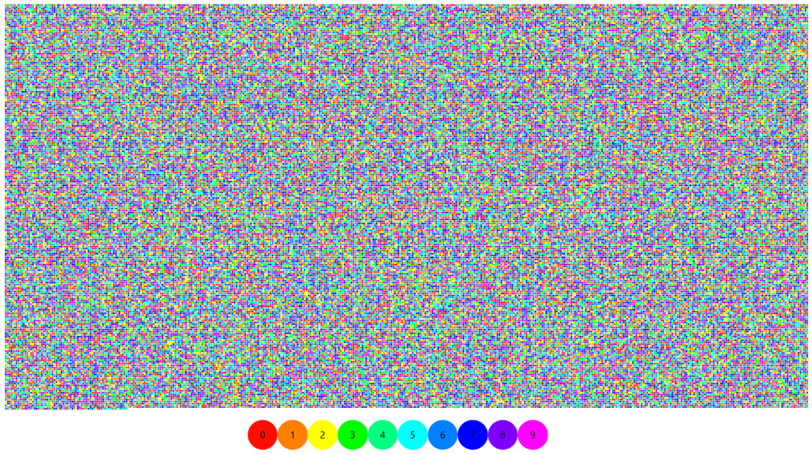

Figura 1: Cómo se ven los primeros 100,000 dígitos de π.

12. Si dos dados están cargados de tal manera que en un dado, el valor 1 aparecerá exactamente el doble de veces que los otros valores, y el otro dado está igualmente cargado hacia el 6, calculen la probabilidad p_s de que un total exactamente igual a s aparecerá en la suma de los dos dados, para $2 \leq s \leq 12$.

13. Algunos dados que fueron cargados como se describe en el problema anterior se lanzaron 144 veces, y se observaron los siguientes valores:

Valor de $s =$	2	3	4	5	6	7	8	9	10	11	12
número observado Y_s	2	6	10	16	18	32	20	13	16	9	2

    Apliquen la prueba $chisq$ a estos valores, usando las probabilidades teóricas para dos dados 'honestos'? ¿La prueba detecta que los dados no son honestos? Si no, explicar porqué no.