

# Módulo Montecarlo Software MES CGE Consultores

René González Leiva  
*aplikadito@gmail.com*

Estudiante Ingeniería Civil Informática  
Universidad Técnica Federico Santa María

Julio, 2008

## 1. Introducción

La simulación Montecarlo básicamente permite obtener muestreos aleatorios de situaciones reales basado en modelos de comportamiento de las variables que participan en cierto problema.

La implementación a realizar en el software MES de la simulación Montecarlo deberá permitir asignar distintas distribuciones de probabilidad a las variables:

- Tiempo entre Fallas (TBF)
- Tiempo de Reparación (TTR)
- Costo de Reparación (CR)

Con estas variables, se calcularán los siguientes indicadores:

- Disponibilidad
- Costo de Mantenimiento
- Costo de la Falta

Los resultados serán probabilísticos, por lo que se utilizan histogramas para presentar la información.

## 2. Distribuciones en Módulo Montecarlo para MES

El módulo Montecarlo para MES implementará las siguientes distribuciones de probabilidad:

- Exponencial  $Exp(\lambda)$
- Weibull  $Weibull(\alpha, \beta)$
- Normal  $N(\mu, \sigma^2)$
- LogNormal  $Ln(\mu, \sigma^2)$
- Triangular  $Tr(a, b, c)$
- Uniforme  $U(a, b)$
- Dirac  $Dirac(p)$

Se definen varios métodos comunes a todas las distribuciones. Se pueden destacar los siguientes métodos:

- Un método que implemente la función de densidad de la distribución  $f(x)$
- Un método que implemente la función de probabilidad acumulada de la distribución  $F(x)$
- Un método que implemente la función inversa de la función de distribución de probabilidad acumulada  $F^{-1}(x)$

### 3. Obtención de números aleatorios

Sea una variable  $X$  que se puede modelar con una exponencial de media 0,007,  $Exp(0,007)$ . Para obtener valores de la variable  $X$  acorde con su comportamiento exponencial, se realiza lo siguiente:

1. Se obtiene un número aleatorio uniformemente distribuido entre 0 y 1 (utilizando la clase **Random** y el método **nextDouble()**)
2. Luego, con ese número, se llama al método que obtiene la función inversa de la función de distribución acumulada.
3. El resultado de la inversa es un número dentro del rango definido por la distribución exponencial asociada.

Por ejemplo, suponiendo que el número aleatorio obtenido fue 0,4. Sabemos que la función de distribución acumulada de una exponencial es

$$F(x) = 1 - e^{-(\lambda x)}$$

La inversa se puede obtener directamente

$$F(x)^{-1} = -\frac{e^{1-y}}{\lambda}$$

donde  $y$  es un número entre 0 y 1. De forma genérica

$$F(x)^{-1} = -\frac{e^{1-U(0,1)}}{\lambda}$$

Ahora, simplemente se reemplazan los valores:  $y = 0,4$ ;  $\lambda = 0,007$

$$F(x)^{-1} = 260,3$$

Con esto, se realiza un mapeo de una función uniforme entre 0 y 1 a una exponencial con media 0.007. Y en general, este procedimiento es similar para toda distribución que posea una función inversa de probabilidad acumulada definida. Este método se conoce como **Algoritmo de la Transformada Inversa** y se puede definir, de forma general, de la siguiente manera:

*Sea  $U$  una variable aleatoria uniforme en  $(0,1)$ . Para cualquier función de distribución continua  $F$ , invertible, la variable aleatoria  $X$  definida como*

$$X = F^{-1}(U)$$

*tiene distribución  $F$  [ $F^{-1}$  se define como el valor de  $x$  tal que  $F(x) = u$ ]*

## 4. Obtención de números aleatorios para Normal y LogNormal

### 4.1. Usando método `nextGaussian()` de la clase `Random`

Dado que no es posible aplicar el método anterior a las distribuciones Normal y LogNormal, se debe utilizar un generador de números aleatorios normalmente distribuido (utilizando la clase `Random` y el método `nextGaussian()`) de manera estándar. Para obtener un número aleatorio respecto a una  $Normal(\mu, \sigma^2)$ , simplemente, se realiza un cambio de variables

$$\begin{aligned}n &= N(0, 1) \\z &= \sigma \cdot n + \mu\end{aligned}$$

donde  $z$  es un número  $N(\mu, \sigma)$ -mente distribuido.

Para obtener un valor aleatorio asociado a una LogNormal con media  $\mu$  y desviación estandar  $\sigma$ , primero, utilizan las siguientes variables auxiliares

$$\begin{aligned}\mu_{aux} &= \lg\left(\frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}}\right) \\ \sigma_{aux} &= \sqrt{\lg\left(1 + \left(\frac{\sigma}{\mu}\right)^2\right)}\end{aligned}$$

Para después realizar el cambio de variables

$$z = e^{\sigma_{aux} \cdot n + \mu_{aux}}$$

donde  $z$  es un número  $Log(\mu, \sigma)$ -mente distribuido. Este es el método que actualmente se está usando.

### 4.2. Método de búsqueda recursiva

Este método no se está utilizando actualmente debido a que su performance no es buena. No obstante, su implementación no ha sido borrada debido a que en el futuro eventualmente podría ser usada. Se adjunta a continuación su explicación

Primero, se define una distribución Normal Estandar,  $N(0,1)$  que sirve de base para los cálculos de las distribuciones de probabilidad acumulada y de la inversa. Las razones se detallan a continuación.

Una distribución Normal no posee una función de distribución acumulada definida. El cálculo para la obtención de esta función se realiza utilizando métodos de integración numérica, o sea, que utilizan la idea básica de Riemann, dividir el dominio de la función en varios intervalos y calcular el área de estos intervalos con polígonos limitados en su altura por los valores de la función, ya sea por exceso o por defecto. El método de Simpson es uno de los mejores métodos de integración numérica que hay. Crea polígonos un poco más complejos a lo largo de la curva y calcula con bastante precisión el área bajo ella. La única dificultad de este método es que necesita de los límites superior e inferior de integración, los cuales no existen para el caso particular de una Normal. Para ello, se utiliza la Normal Estandar, para la cual se conocen límites superiores e inferiores que son factibles de utilizar. En este caso, se utiliza como límite inferior  $-5$  y límite superior  $5$ , dado que las probabilidades de que los números estén fuera del intervalo  $[-5,5]$  es ínfima. Por lo tanto, se utilizan estos límites para decir que una Normal Estandar tiene un dominio que sólo existe dentro del intervalo ya definido. Ahora, es posible utilizar el método de Simpson para integrar una Normal Estandar.

Si se desea obtener la función de probabilidad acumulada para una  $Normal(\mu, \sigma^2)$  para un valor  $x$  (es decir, se desea saber cuánto vale la integral desde  $-\infty$  hasta el punto  $x$ ) se puede utilizar el siguiente cambio de variable:

$$z = \frac{x - \mu}{\sigma}$$

donde  $z$  es un valor mapeado para el dominio de una Normal Estandar. Luego, se calcula la integral de una Normal Estandar desde  $-\infty$  (que en nuestro caso particular hemos definido como  $-5$ ) hasta el punto  $z$  usando el método de Simpson.

De forma idéntica, para una LogNormal se utiliza:

$$z = \frac{\ln(x) - \mu}{\sigma^2}$$

y se calcula la integral para la Normal Estandar usando  $z$ .

Finalmente, es posible realizar el cálculo de la función inversa de la distribución acumulada usando un algoritmo recursivo. Para ello, se debe realizar un mapeo inverso. Primero, supongamos que se tiene un número  $\mathbf{p}$  aleatorio uniformemente distribuido entre 0 y 1. Lo que deseamos saber es con que valor  $\mathbf{z}$  en la función de distribución acumulada de una Normal Estandar es posible obtener el valor  $\mathbf{p}$ . En otras palabras, encontrar el valor  $\mathbf{z}$  tal que la integral entre  $-\infty$  y  $\mathbf{z}$  para una Normal Estandar sea igual a  $\mathbf{p}$ .

Este valor  $\mathbf{z}$  se obtiene realizando una búsqueda binaria entre los límites de la Normal Estandar anteriormente definidos de la siguiente manera:

1. Se comienza con límites definidos (  $[-5,5]$  )
2. Se evalúa en la mitad de los límites actuales (  $F(0)$  )
3. Si el valor objetivo  $\mathbf{p}$  es mayor que la evaluación en el punto medio, se corre el límite inferior hasta el punto medio y se vuelve a evaluar en la mitad de los nuevos límites.
4. Si el valor objetivo  $\mathbf{p}$  es menor que la evaluación en el punto medio, se corre el límite superior hasta el punto medio y se vuelve a evaluar en la mitad de los nuevos límites.
5. El algoritmo para cuando se la diferencia entre el punto evaluado y la meta  $\mathbf{p}$  es menor a cierta precisión.

Esto es análogo al proceso de búsqueda binaria de, por ejemplo, un nombre en un directorio telefónico, buscar iterativamente en las mitades de las mitades.

De esta forma, se obtiene un valor  $\mathbf{z}$  que cumple las condiciones ya mencionadas. Este valor se mapea para el dominio de una  $Normal(\mu, \sigma^2)$  de la siguiente manera:

$$x = \sigma * z + \mu$$

donde  $\mathbf{x}$  es el valor mapeado para una  $Normal(\mu, \sigma^2)$  respecto de  $\mathbf{z}$  en una Normal Estandar. Para una  $LogNormal(\mu, \sigma^2)$  el cambio de variable es:

$$x = e^{\sigma * z + \mu}$$

donde  $\mathbf{x}$  es el valor mapeado para una  $LogNormal(\mu, \sigma^2)$  respecto de  $\mathbf{z}$  en una Normal Estandar.

A pesar de tener definido y resuelto el problema de la obtención de inversas, es necesario definir cotas inferiores y superiores para las distribuciones, debido al problema computacional de graficar las funciones y además del requerimiento para el módulo de que las distribuciones puedan ser truncadas por cotas definidas por el usuario. Se deben definir cotas superiores e inferiores por defecto. Éstas, arbitrariamente, se definen de la siguiente manera:

- **Cota Inferior:** un valor  $\mathbf{x\_inf}$  tal que la probabilidad desde  $-\infty$  hasta aquel punto sea **0,1 %**. O sea, la función inversa de probabilidad acumulada de la distribución para el valor **0,001**.
- **Cota Superior:** el valor  $\mathbf{x\_sup}$  tal que la probabilidad desde  $-\infty$  hasta aquel punto sea **99,9 %**. O sea, la función inversa de probabilidad acumulada de la distribución para el valor **0.999**.