SIMULACION

Simulación: es una herramienta para la toma de decisiones. Es el proceso de diseñar un modelo de un sistema real y llevar a cabo experiencias con el mismo; con la finalidad de comprender el comportamiento del sistema o evaluar nuevas estrategias (dentro de los límites impuestos por un criterio o conjunto de ellos), para el funcionamiento del sistema.
La simulación se utiliza porque:

- Ensayar sobre sistemas reales puede llevar a la destrucción de los mismos.
- Ensayar sobre sistemas reales (prototipos) puede ser muy costoso.
- En los estudios de determinados sistemas puede interesar alterar las escalas del tiempo.

La simulación es una técnica que ayuda a construir un modelo de una situación real y realizar experimentos.
El objetivo de la simulación es la obtención de información decisoria que permita mejorar la predicción implícita en toda decisión.

Sistema: es un conjunto de elementos que ordenadamente relacionados entre sí contribuyen a un determinado objetivo.

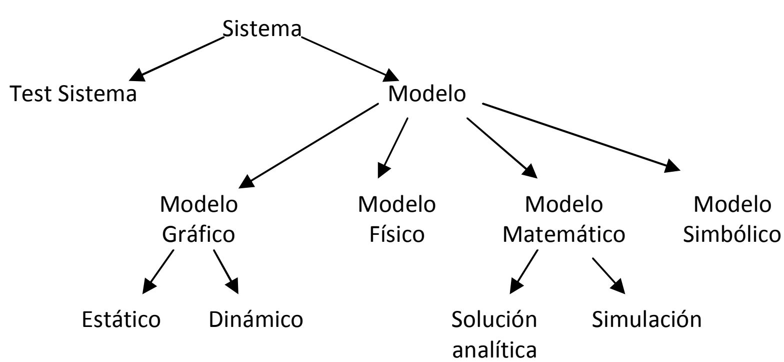
Modelo: es una representación del sistema que puede ser experimentado y manipulado con el fin último de estudiar al sistema. Es una representación que pretende ser el sistema en estudio.
Los modelos se utilizan esencialmente para hacer tres cosas:

1. Estudiar el sistema existente sin molestar su operación.
2. Estudiar sistemas existentes sin destruirlos.
3. Estudiar sistemas futuros inexistentes.

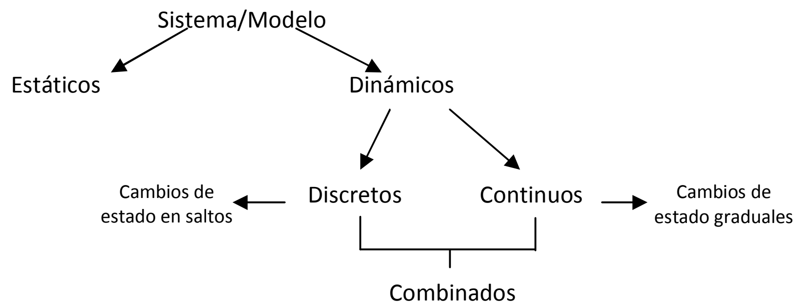
Todo proceso de modelización implica:

1. Identificación de las entidades principales y de sus atributos característicos.
2. Identificación y representación de las reglas que gobiernan el sistema que se quiere simular.
3. Captación de la naturaleza de las interacciones lógicas del sistema que se modela.
4. Verificación de que las reglas incorporadas al modelo son representaciones validas de las del sistema que se modela.
5. Representación del comportamiento aleatorio.

$\Rightarrow$ **Utilizamos un modelo e invertimos en simulación para ahorrar dinero.**



Solución Analítica: establece un modelo determinista, donde no existe ningún tipo de probabilidad asociada. Plantea una seguridad absoluta respecto al resultado a obtener. Se obtiene una solución única. Ejemplos de una solución analítica es una ecuación matemática o una fórmula física.
Simulación: los resultados no tienen seguridad absoluta sino que, utilizando las herramientas de probabilidad, intenta determinar el porcentaje de seguridad del resultado.



Sistemas estáticos: se mantienen estables a través del tiempo. Ej: Montecarlo.
Sistemas dinámicos: el estado en los sistemas varía con el tiempo.

Elementos que componen un sistema

1. **Entidad:** es la representación de los flujos de entrada a un sistema, es el elemento responsable de que el estado de un sistema cambie.
2. **Estado del sistema:** es el conjunto de algunas variables prescriptas en un sistema en algún punto del tiempo. Es la foto del sistema en un momento dado. Es la condición que guarda el sistema bajo estudio en un momento dado.
3. **Evento:** es una situación que genera un cambio actual del sistema.

    Actuales: son aquellos que están sucediendo en el sistema en un momento dado.
    Futuros: son cambios que se producirán en el estado del sistema en algún instante futuro de acuerdo con una programación específica.

4. **Atributo:** es una característica de una entidad.
5. **Variables:** son condiciones cuyos valores se crean o modifican por medio de ecuaciones matemáticas o relaciones lógicas.
6. **Reloj de simulación:** es el contador de tiempo de la simulación.

    Absoluto: parte de cero y termina en un tiempo total de simulación definido.
    Relativo: solo considera el lapso de tiempo que transcurre entre dos eventos.

7. **Recursos:** dispositivos necesarios para llevar a cabo una operación.

Usos y limitaciones de la simulación

Uso cuando:

- No existe una formulación matemática completa del problema.
- Existen los métodos analíticos pero, las hipótesis simplificadoras necesarias para su aplicación, desvirtúan las soluciones obtenidas y su interpretación.
- Los métodos analíticos existen y en teoría están disponibles, pero son complejos.
- La simulación constituye la mejor alternativa por la dificultad de realizar experiencias en el contexto real.

⇒ **Se usa la simulación cuando no existe otra técnica para encarar la resolución del problema.**

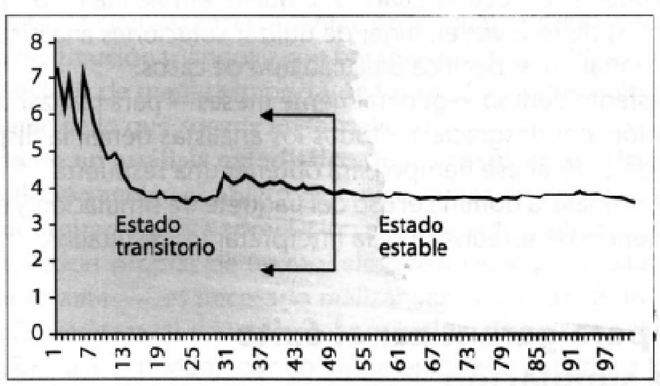
Limitaciones:

- Costo en horas de desarrollo y de computadoras (procesamiento).
- Dificultad en la validación del modelo y de los resultados.
- Recolección, análisis e interpretación de los resultados suele requerir personal con conocimientos estadísticos.
- La aceptación de los resultados requiere un elevado conocimiento del modelo empleado.

Replica o Corrida

Es necesario efectuar más de una réplica del modelo que se esté analizando, con la finalidad de obtener estadísticas de intervalo que nos den una mejor ubicación del verdadero valor de la variable, bajo los diferentes escenarios que se presentan al modificar los números pseudoaleatorios en cada oportunidad.
Se presentan dos etapas en la simulación: un **estado transitorio** y un **estado estable**. El primero se presenta al principio de la simulación.
En el estado estable los valores de las variables de decisión permanecen muy estables, presentando solo variaciones poco significativas. En este momento las decisiones que se tomen serán mucho más confiables.
Sin embargo, no todas las variables convergen al estado estable con la misma rapidez.
Otro factor importante para decidir el tiempo de simulación es el costo de la corrida.

velocidad



tiempo

Ventajas y Desventajas de la simulación

Ventajas:

1. Es muy buena herramienta para conocer el impacto de los cambios en los procesos sin necesidad de llevarlos a cabo en la realidad.
2. Mejora el conocimiento del proceso actual al permitir que el analista vea cómo se comporta el modelo generado bajo diferentes escenarios.
3. Puede utilizarse como medio de capacitación para la toma de decisiones.
4. Es más económico realizar un estudio de simulación que hacer muchos cambios en los procesos reales.
5. Permite probar varios escenarios en busca de las mejores condiciones de trabajo de los procesos que se simulan.
6. En problemas de gran complejidad, la simulación permite generar una buena solución.
7. La simulación es una ventaja frente a los modelos de análisis altamente complejos que no pueden resolverse de forma analítica.
8. Los modelos de simulación son útiles en la realidad donde no debemos perturbar los sistemas.
9. Planteados los modelos, por medio de computadoras, podemos replicarlas cuantas veces sea necesario.

Desventajas:

1. La simulación no es una herramienta de optimización.
2. La simulación es costosa.
3. Se requiere bastante tiempo para realizar un buen estudio de simulación, pues “las cosas no vienen dadas”.
4. No es posible asegurar que el modelo sea válido, debido a la subjetividad del analista que lo realiza.

Elementos claves a tener en cuenta:

Independientemente de los beneficios que conlleva la simulación, es imposible garantizar que un modelo tendrá éxito.
Algunas causas por las que un modelo de simulación podría no tener los resultados que se desean son:

- Tamaño insuficiente de la corrida.
- Variables de respuestas mal definidas.
- Errores al establecer las relaciones entre las variables aleatorias.
- Errores al determinar el tipo de distribución asociado a las variables aleatorias del modelo.
- Falta de un análisis estadístico de los resultados.
- Uso incorrecto de la información obtenida.
- Falta o exceso de detalle del modelo.

El proceso de simulación incluye lo siguiente:

Se aísla lo que se quiere estudiar	→	se define el sistema
Se lo representa para estudiarlo (se tiene en cuenta lo relevante)	→	se construye el modelo
Se realizan ensayos en el modelo	→	se simula
Se sacan conclusiones	→	se infiere lo que va a pasar
Se estudian las conclusiones	→	se aconseja la mejor

Pasos de la simulación:

1. **Definición del sistema bajo estudio:** es necesario conocer el sistema a modelar. Es conveniente definir con claridad las variables de decisión del modelo, determinar las interacciones entre estas y establecer con precisión los alcances y limitaciones que aquel podría llegar a tener. Se debe tener en cuenta cuál es el objetivo de la simulación.
2. **Generación de un modelo de simulación base:** no es preciso que este modelo sea demasiado detallado. Podemos ver qué ecuaciones plantear y sus relaciones, y luego comenzar a testear las ecuaciones planteadas frente a la realidad.
3. **Recolección y análisis de datos:** es la etapa más larga. Se debe determinar qué información es útil para la determinación de las distribuciones de probabilidad asociadas a cada una de las variables aleatorias necesarias para la simulación.
4. **Generación del modelo preliminar:** se integra información obtenida a partir del análisis de datos, los supuestos del modelo y todos los datos que se requieran para obtener un modelo lo más cercano a la realidad del problema bajo estudio.
5. **Verificación del modelo:** para comprobar la propiedad de la programación del modelo, y comprobar que todos los parámetros usados en la simulación funcionan correctamente.
6. **Validación del modelo:** consiste en realizar una serie de pruebas al mismo, utilizando información de entrada real para observar su comportamiento y analizar sus resultados.
7. **Generación del modelo final.**
8. **Determinación de los escenarios para el análisis.**
9. **Análisis de sensibilidad:** realizar pruebas estadísticas para comparar los distintos escenarios. Se pueden realizar más replicas de un escenario determinado y también se puede alargar o extender el tiempo de simulación.
10. **Documentación del modelo, sugerencias y conclusiones:** implica plantear la toma de decisiones, recomendaciones de acuerdo a los resultados y elaborar informe si es necesario.

SIMULACION ESTATICA (MONTECARLO)

Teorema de Bernoulli: cuando no conocemos las probabilidades asociadas a los sucesos bajo análisis (caso más común en la realidad) y "la experiencia es grande", podemos tomar la frecuencia relativa del acontecimiento como valor aproximado de las probabilidades.

Utilizamos Montecarlo cuando se tienen variables de comportamiento aleatorios. Está basada en el muestreo sistemático de variables aleatorias.

Montecarlo surge por dos motivos:

- Falta de tiempo y recursos para determinar la distribución de probabilidad.
- Es un método rápido. Útil porque utiliza datos históricos acerca de lo que interesa.

Posibilita estas ventajas:

- La precisión dada se puede incrementar aumentando el número de experiencias.
- Mejora el procesamiento de los datos.
- Repeticiones en tiempo corto.
- Facilidad de uso.

Los componentes primarios de un método de simulación Montecarlo incluyen lo siguiente:

- Funciones de distribución de probabilidad generados a partir de las frecuencias relativas.
- Generador de números aleatorios.
- Regla de muestreo.
- Anotar (o registrar).
- Estimación del error.
- Técnicas de reducción de la varianza.
- Paralelización y vectorización. Por ejemplo, en aplicaciones con muchas variables se estudia trabajar con varios procesadores paralelos para realizar la simulación.

NUMEROS PSEUDOALEATORIOS

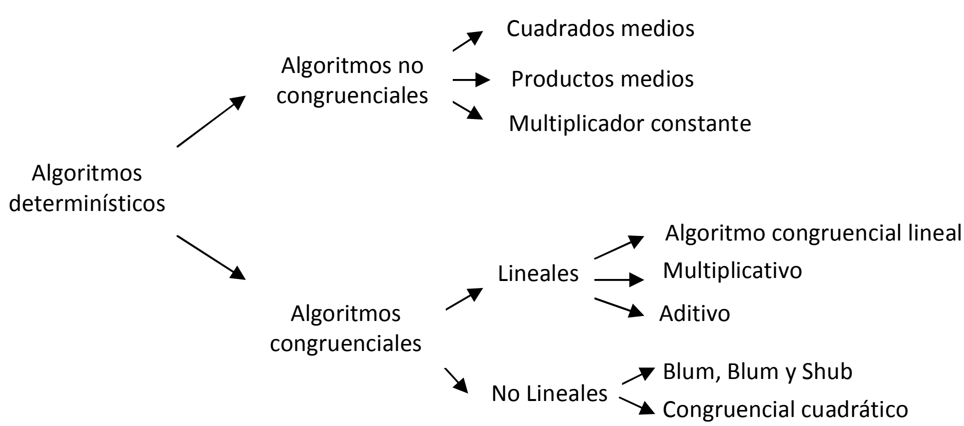
Generación de números pseudoaleatorios

Para poder realizar una simulación que incluya variabilidad dentro de sus eventos, es preciso generar una serie de números que sean aleatorios por sí mismos, y que su aleatoriedad se extrapole al modelo de simulación que se está construyendo.
Para realizar una simulación se requieren números aleatorios en el intervalo (0,1), a los cuales se hará referencia como r_i, es decir una secuencia r_i=(r_1 ,r_2, … , r_n) que contiene **n** números, todos ellos diferentes, **n** recibe el nombre de **periodo** o **ciclo de vida** del generador que creó la secuencia r_i.
Debido a que no es posible generar números realmente aleatorios, consideramos los r_i como números pseudoaleatorios, generados por medio de algoritmos determinísticos que requieren parámetros de arranque.
Para simular el comportamiento de 1 o más variables aleatorias es necesario contar con un conjunto suficientemente grande de r_i.
Los resultados no pueden basarse en una sola simulación del sistema, por el contrario es necesario realizar varias replicas de la misma, corriendo cada una de ellas con números pseudoaleatorios diferentes.
El conjunto r_i debe ser sometido a una variedad de pruebas para verificar si los números que lo conforman son realmente uniformes e independientes.
Un conjunto r_i debe seguir una distribución uniforme continua, la cual está definida por:

$$f(r)=\begin{cases} 1 & 0\leq r \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Problemas:

1. Que los números del conjunto r_i no estén uniformemente distribuidos, es decir, que haya demasiados r_i en un subintervalo y en otros muy pocos o ninguno.
2. Que los números r_i generados sean discretos en lugar de continuos.
3. Que la media del conjunto sea muy alta o muy baja, es decir, que esté por arriba o por debajo de ½.
4. Que la varianza del conjunto sea muy alta o muy baja, es decir, que se localice por arriba o por debajo de 1/12.



Algoritmo de cuadrados medios

Requiere un número entero detonador (llamado semilla) con D dígitos, el cual es elevado al cuadrado para seleccionar del resultado los D dígitos del centro, el primer número r_i se determina anteponiendo simplemente el "0." a esos dígitos. Para obtener el segundo numero r_i se sigue el mismo procedimiento, solo que ahora se elevan al cuadrado los D dígitos del centro que se seleccionaron para obtener el primer r_i.

Pasos:

1. Seleccionar la semilla (X_0) con D dígitos (D› 3).
2. Sea Y_0 igual al resultado de elevar X_0 al cuadrado, sea X_1 igual a los D dígitos del centro, y sea r_1= 0.D dígitos del centro.
3. Sea Y_i igual al resultado de elevar X_i al cuadrado, sea X_{i+1} igual a los D dígitos del centro, y sea r_i= 0.D dígitos del centro para toda i= 1, 2, 3,…, n.
4. Repetir el paso 3 hasta obtener los n números r_i deseados.

Este algoritmo es incapaz de generar una secuencia de r_i con período de vida n grande.

Algoritmo de productos medios

Requiere dos semillas, ambas con D dígitos. Las semillas se multiplican y del producto se seleccionan los D dígitos del centro, los cuales formarán el primer número pseudoaleatorio r_1 = 0.D dígitos. Después se elimina una semilla, y la otra se multiplica por el primer número de D dígitos, para luego seleccionar del producto los D dígitos que conformarán un segundo numero r_i.

Pasos:

1. Seleccionar una semilla (X_0) con D dígitos (D› 3).
2. Seleccionar una semilla (X_1) con D dígitos (D› 3).
3. Sea Y_0 =(X_0*X_1) , sea X_2 igual a los D dígitos del centro, y sea r_1= 0.D dígitos del centro.
4. Sea Y_i =(X_i*X_{i+1}), sea X_{i+2} igual a los D dígitos del centro, y sea r_{i+1} = 0.D dígitos del centro para toda i= 1, 2, 3,…, n.
5. Repetir el paso 4 hasta obtener los n números r_i deseados.

Algoritmo de multiplicador constante

Pasos:

1. Seleccionar una semilla (X_0) con D dígitos (D› 3).
2. Seleccionar una constante (a) con D dígitos (D› 3).
3. Sea Y_0 =(a*X_0) , sea X_1 igual a los D dígitos del centro, y sea r_1= 0.D dígitos del centro.
4. Sea Y_i =(a*X_i), sea X_{i+1} igual a los D dígitos del centro, y sea r_{i+1} = 0.D dígitos del centro para toda i= 1, 2, 3,…, n.
5. Repetir el paso 4 hasta obtener los n números r_i deseados.

Algoritmo lineal

Genera una secuencia de números enteros por medio de la siguiente ecuación recursiva:

$$X_{i+1} = (a.X_i + c) \bmod (m) \qquad i=0, 1, 2,..., n$$

Donde $\mathbf{X_0}$ es la semilla, **a** es la constante multiplicativa, **c** es una constante aditiva y **m** es el módulo; $X_0 > 0$, $a > 0$, $c > 0$ y $m > 0$ deben ser números enteros.
La ecuación recursiva del algoritmo congruencial lineal genera una secuencia de números enteros S = {0, 1, 2, …, m-1}, y para obtener los números pseudoaleatorios en el intervalo (0,1) se requiere la siguiente ecuación:

$$r_i = \frac{X_i}{m-1} \qquad i = 1,2, \dots, n$$

Condiciones:

- $m = 2^g$
- $a = 1+4k$
- k debe ser entero
- c relativamente primo a m
- g debe ser entero

Bajo estas condiciones se obtiene un período de vida máximo: $N = m = 2^g$.

Algoritmo congruencial multiplicativo

Surge del anterior, cuando c = 0. Entonces, la ecuación recursiva es:

$$X_{i+1} = (a.X_i) \bmod (m) \qquad i=0, 1, 2,..., n$$

La ventaja que tiene es que implica una operación menos a realizar.

Condiciones:

- $m = 2^g$
- $a = 3+8k$ o $a = 5+8k$
- k debe ser entero
- X_0 debe ser un número impar
- g debe ser entero

$\Rightarrow N = m/4 = 2^{g-2}$.

Algoritmo congruencial aditivo

Este requiere una secuencia previa de n números enteros $X_1, X_2, \dots, X_n$ para generar una nueva secuencia de números enteros que empiezan en $X_{n+1}, X_{n+2}, \dots$, . Su ecuación recursiva es:

$$X_i = (X_{i-1} + X_{i-n}) \bmod (m) \qquad i= n+1, n+2,..., N$$

Los números pueden ser generados mediante la ecuación:

$$r_i = \frac{X_i}{m-1} \qquad i = 1,2, \dots, n$$

Algoritmo congruencial cuadrático

Tiene la siguiente ecuación recursiva:

$$X_{i+1} = (a.X_i^2 + b.X_i + c) \bmod (m) \qquad i=0, 1, 2,..., N$$

En este caso los números r_i pueden ser generados con la ecuación:

$$r_i = \frac{X_i}{m-1} \qquad i = 1,2, \dots, n$$

Condiciones:

- $m = 2^g$
- a debe ser un número par
- c debe ser un número impar
- g debe ser entero
- (b-1) mod 4 = 1

$\Rightarrow N = m$

Algoritmo de Blum, Blum y Shub

Si en el algoritmo congruencial cuadrático a=1, b=0 y c=0, entonces:

$$X_{i+1} = (X_i^2) \bmod (m) \qquad i=0, 1, 2,..., N$$

Propiedades de los números aleatorios

- **Media de los aleatorios entre 0 y 1:** en vista de que estos números deben tener la misma probabilidad de presentarse, es preciso que su comportamiento muestre una distribución de probabilidad uniforme continua, con límite inferior cero y límite superior uno. La función de densidad de una distribución uniforme es la siguiente:

$$f(x) = \frac{1}{b-a} \qquad a \leq x \leq b \qquad \text{en este caso a=0 y b=1}$$

$$E(x) = \int_a^b \frac{1}{b-a}(x)dx = \frac{x^2}{2(b-a)} \rightarrow E(x) = 1/2$$

El valor esperado es μ=0,5.

- **Varianza de los números aleatorios:**

$$V(x) = \sigma^2 = E(x^2) - \mu^2$$

$$E(x^2) = \int_a^b \frac{1}{b-a}(x^2)dx = \frac{x^3}{3(b-a)} \quad \rightarrow E(x^2) = 1/3$$

$$V(x) = E(x^2) - \mu^2 = \tfrac{1}{3} - \left(\tfrac{1}{2}\right)^2 = \tfrac{1}{12} \qquad \rightarrow \sigma^2 = 1/12$$

- **Independencia:** implica que los números aleatorios no deben tener correlación entre sí, es decir, deben ser independientes de manera que puedan dispersarse uniformemente dentro de todo el espectro de valores posibles.

PRUEBAS ESTADISTICAS PARA LOS NUMEROS PSEUDOALEATORIOS

Prueba de las medias

Una de las propiedades que deben cumplir los números del conjunto r_i, es que el valor esperado sea igual a 0,5.

Hipótesis

- $H_0 : \mu_{ri} = 0{,}5$
- $H_1 : \mu_{ri} \neq 0{,}5$

La prueba de las medias consiste en determinar el promedio de los n números que contiene el conjunto r_i, mediante la siguiente ecuación:

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$$

Posteriormente se calculan los límites de aceptación inferior y superior con las ecuaciones siguientes:

$$Li_{\bar{r}} = \tfrac{1}{2} - z_{\alpha/2}\left(\tfrac{1}{\sqrt{12n}}\right) \qquad\qquad Ls_{\bar{r}} = \tfrac{1}{2} + z_{\alpha/2}\left(\tfrac{1}{\sqrt{12n}}\right)$$

Si el valor de $\bar{r}$ se encuentra entre los límites de aceptación, concluimos que no se puede rechazar que el conjunto r_i tiene un valor esperado de 0,5 con un nivel de aceptación de 1-α. En caso contrario se rechaza que el conjunto r_i tiene un valor esperado de 0,5.
Para el cálculo de los límites de aceptación se utiliza el estadístico $z_{\alpha/2}$, el cual se determina por medio de la tabla de distribución normal estándar.

Prueba de la varianza

Otra de las propiedades que debe satisfacer el conjunto r_i, es que sus números tengan una varianza de 1/12. Si no pasa esta prueba puede ser que los números estén demasiados alrededor del 0,5.

Hipótesis

- $H_0 : \sigma^2_{ri} = 1/12$
- $H_1 : \sigma^2_{ri} \neq 1/12$

La prueba de la varianza consiste en determinar la varianza de los n números que contiene el conjunto r_i, mediante la siguiente ecuación:

$$V(\bar{r}) = \frac{\sum_{i=1}^{n}(r_i - \bar{r})^2}{n-1}$$

$$Li_{V(\bar{r})} = \frac{X^2_{\propto/2,n-1}}{12(n-1)} \qquad Ls_{V(\bar{r})} = \frac{X^2_{1-\propto/2,n-1}}{12(n-1)}$$

Si el valor de $V(\bar{r})$ se encuentra entre los límites de aceptación, decimos que no se puede rechazar que el conjunto r_i tiene una varianza de 1/12, con un nivel de aceptación de 1-α, de lo contrario se rechaza que el conjunto r_i tiene una varianza de 1/12.
Para el cálculo de los límites de aceptación se utiliza el estadístico $X^2_{\propto/2,n-1}$, el cual se determina por medio de la tabla de Chi-Cuadrada.

> **IMPORTANTE:** en la calculadora, estando en modo estadístico (SD), la opción xσn-1 se debe elevar al cuadrado para obtener la varianza: $(x\sigma n\text{-}1)^2 = V(r)$ OJO!!!!!

Prueba de uniformidad

Hipótesis

- $H_0 : r_i \sim U(0,1)$
- $H_1 : r_i \neq U(0,1)$

Prueba Chi-Cuadrada

Busca determinar si los números del conjunto r_i se distribuyen uniformemente en el intervalo (0,1). Para llevar a cabo esta prueba es necesario dividir el intervalo (0,1) en m subintervalos, en donde es recomendable $m = \sqrt{n}$. Posteriormente se clasifica cada número pseudoaleatorio del conjunto r_i en los m intervalos. A la cantidad de números de r_i que se espera encontrar en cada intervalo se llama frecuencia esperada (E_i), teóricamente la $E_i = n/m$. A partir de los valores O_i y E_i se determina el estadístico X^2_o mediante la ecuación:

$$X_0^2 = \sum_{i=1}^{m} \frac{(E_i - O_i)}{E_i}$$

Si el valor del estadístico X^2_o es menor al valor de tablas $X^2_{\propto,m-1}$ entonces, no se puede rechazar que el conjunto r_i sigue una distribución uniforme. En caso contrario, se rechaza que r_i sigue una distribución uniforme.

Pasos:

1. Se subdivide el intervalo (0,1) en m subintervalos, siendo conveniente que $m = \sqrt{n}$
2. Clasificar cada número r_i en cada subintervalo.
3. Determinar la frecuencia observada O_i para cada subintervalo.
4. Determinar la frecuencia esperada E_i para cada subintervalo: $E_i = n/m$.
5. Determinar el estadístico:

$$X_0^2 = \sum_{i=1}^{m} \frac{(E_i - O_i)^2}{E_i}$$

6. Si $X_0^2 < X^2_{\propto,m-1}$ $\Rightarrow$ no se puede rechazar H_0

Se utiliza el estadístico $X^2_{\propto,m-1}$, el cual se determina por medio de la tabla de Chi-Cuadrada.

Prueba de Kolmogorov-Smirnov

Es recomendable aplicarla en conjuntos r_i pequeños, por ejemplo con n<20.

Pruebas de independencia

Demuestra que no existe correlación entre números aleatorios.

Hipótesis

- H_0 : los números del conjunto r_i son independientes
- H_1 : los números del conjunto r_i no son independientes

Pruebas de independencia

- Prueba de las corridas arriba y abajo
- Prueba de las corridas arriba y debajo de las medias
- Prueba de póker
- Prueba de series
- Prueba de huecos

Prueba de series

Consiste en comparar los números con el propósito de corroborar la independencia entre números consecutivos.
Se inicia creando una gráfica de dispersión entre números consecutivos (r_i, r_{i+1}); posteriormente se divide la gráfica en m casillas, siendo m el valor entero más cercano a $\sqrt{n}$ que permita formar de preferencia una matriz cuadrada.
Enseguida se determina la frecuencia observada O_i, contabilizando los números de puntos de cada casilla y su correspondiente frecuencia esperada E_i, de acuerdo con E_i = (n-1)/m, donde n-1 es el número total de pares ordenados o puntos de la gráfica. Se procede entonces a calcular el error o estadístico; finalmente si el valor del error es menor o igual al estadístico de tablas $X^2_{\propto,m-1}$, no podemos rechazar la hipótesis de independencia entre números consecutivos.

Pasos:

1. Crear una gráfica de dispersión entre los números consecutivos (r_i, r_{i+1})
2. Dividir la gráfica en m casillas, $m = \sqrt{n}$, buscando construir una matriz cuadrada.
3. Determinar la frecuencia observada O_i , contabilizando el número de puntos de cada casilla.
4. Determinar el valor de la frecuencia esperada E_i = n-1/m, donde n-1 es el número total de pares ordenados.
5. Determinar el estadístico:
$$X^2_0 = \sum_{i=1}^{m} \frac{(E_i - O_i)^2}{E_i}$$
6. Si $X^2_0 < X^2_{\propto,m-1}$ $\Rightarrow$ no se puede rechazar H_0

IMPORTANTE: tener en cuenta que

- 1-α = nivel de aceptación. Ej: 0,95
- α = nivel de rechazo. Ej: 0.05

VARIABLES ALEATORIAS

Las variables aleatorias son aquellas que tienen un comportamiento probabilístico en la realidad. Usamos la variable aleatoria para generar una función que permita determinar los valores aleatorios para una determinada función de distribución. En cambio los números aleatorios son el alimento para el generador y los usamos para asegurarnos que la respuesta sea aleatoria, pues justamente ese ingreso es aleatorio.

Reglas de distribución de probabilidad

- **Variables aleatorias discretas**

  1. $P(x) \geq 0$
  2. $\sum_{i=1}^{\infty} p_i = 1$
  3. $P(a \leq x \leq b) = \sum_{i=a}^{b} p_i = \ p_a + \cdots + p_b$

  Ej: función uniforme discreta, la de Bernoulli, hipergeométrica, Poisson y binomial.

- **Variables aleatorias continuas**

  1. $P(x) \geq 0$
  2. $P(x = a) = 0$
  3. $\int_{-\infty}^{\infty} f(x) = 1$
  4. $P(a \leq x \leq b) = P(a < x < b) = \ \int_a^b f(x)$

  Ej: función uniforme continua, exponencial, normal, de Weibull, Chi-Cuadrada y la de Erlang.

Determinación del tipo de distribución de un conjunto de datos

La distribución de probabilidad de los datos históricos puede determinarse mediante las pruebas: Chi-Cuadrada, de Kolmogorov-Smirnov y de Anderson Darling.

Prueba Chi-Cuadrada

Se trata de una prueba de hipótesis a partir de datos, basada en el cálculo de un valor llamado **estadístico de prueba**, al cual suele comparársele con un valor conocido como **valor crítico**, mismo que se obtiene generalmente de tablas estadísticas.

Pasos:

1. Obtener al menos 30 datos de la variable aleatoria a analizar.
2. Calcular la media y la varianza de los datos.
3. Crear un histograma de $m = \sqrt{n}$ intervalos, y obtener la frecuencia observada en cada intervalo O_i.
4. Establecer explícitamente la hipótesis nula, proponiendo una distribución de probabilidad que se ajusta a la forma del histograma.
5. Calcular la frecuencia esperada E_i, a partir de la función de probabilidad propuesta.
6. Calcular el estadístico de prueba:
$$C = \sum_{i=1}^{m} \frac{(E_i - O_i)^2}{E_i}$$
7. Definir un nivel de significancia de la prueba, α, y determinar el valor critico de la prueba: $X^2_{\propto,m-k-1}$
8. Comparar el estadístico de prueba con el valor crítico. Si C< $X^2_{\propto,m-k-1}$ entonces no se puede rechazar la hipótesis nula.

Prueba de Kolmogorov-Smirnov

Su limitante es que se puede aplicar solo al análisis de variables continuas. Los primeros 3 pasos son igual a la anterior prueba.

4. Calcular la probabilidad observada en cada intervalo $PO_i = O_i/n$, esto es dividir la frecuencia observada O_i por el número total de datos n.
5. Acumular las probabilidades PO_i para obtener la probabilidad observada hasta el i-ésimo intervalo, POA_i.
6. Establecer explícitamente la hipótesis nula, proponiendo una distribución de probabilidad que se ajuste a la forma del histograma.
7. Calcular la probabilidad esperada acumulada para cada intervalo, PEA_i, a partir de la función de probabilidad propuesta.
8. Calcular el estadístico de prueba:
$$C = máx[PEA_i - POA_i] \qquad i = 1,2,\dots,k,\dots,n$$
9. Definir el nivel de significancia de la prueba α, y determinar el valor critico $D_{\alpha,n}$.
10. Comparar el estadístico de prueba con el valor crítico. Si C< $D_{\alpha,n}$ entonces no se puede rechazar la hipótesis nula.

GENERACION DE VARIABLES ALEATORIAS

Métodos
- De la transformada inversa
- De convolución
- De composición
- De la transformada directa
- De aceptación y rechazo

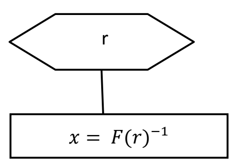
Método de la transformada inversa

Puede utilizarse para simular variables aleatorias continuas, lo cual se logra mediante la función acumulada F(x) y la generación de números pseudoaleatorios r_i ~ U(0,1).

Pasos:

1. Definir la función de densidad f(x) que representa la variable a modelar.
2. Calcular la función acumulada F(x).
3. Despejar la variable aleatoria x y obtener la función acumulada inversa $F(x)^{-1}$.
4. Generar las variables aleatorias x, sustituyendo valores con números pseudoaleatorios entre 0 y 1, en la función acumulada inversa.

$$x = F(r)^{-1} \quad \rightarrow donde\ r\ es\ un\ numero\ aleatorio$$



- **Distribución uniforme**

$$f(x) = \frac{1}{b-a} \qquad a \le x \le b \quad \rightarrow F(x) = \int_0^x \frac{1}{b-a} dx = \frac{x-a}{(b-a)}$$

Igualando a la función acumulada F(x) con el número r_i, y despejando x:

$$x_i = a + (b-a)F(x_i) \quad \rightarrow \quad x_i = a + (b-a)r_i$$

- **Distribución exponencial**

$$f(x) = \lambda\, e^{-\lambda x} \quad para\ x \ge 0 \qquad \rightarrow F(x) = \int_0^x \lambda\, e^{-\lambda x}\, dx = 1 - e^{-\lambda x}$$

$$x_i = \frac{-1}{\lambda} \ln\,(1 - F(x_i)) \quad \rightarrow \quad x_i = \frac{-1}{\lambda} \ln\,(1 - r_i)$$

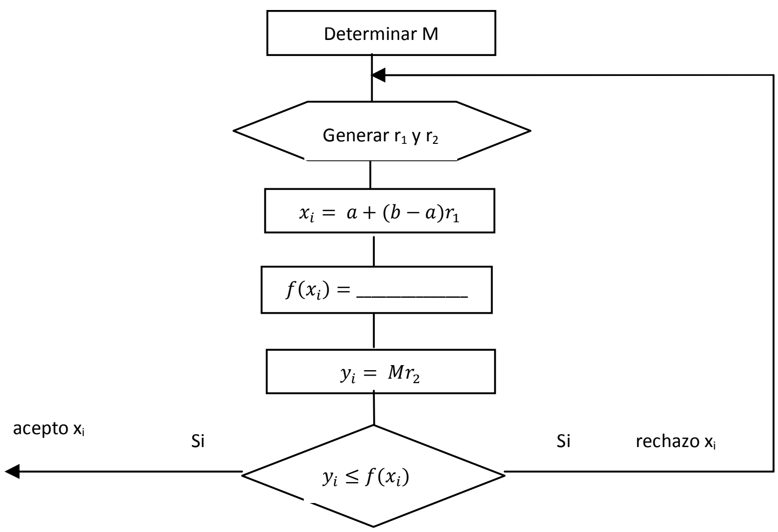
Método del rechazo

Si consideramos que la distribución normal viene dada por: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/x\left(\frac{x-\sigma}{t}\right)^2}$

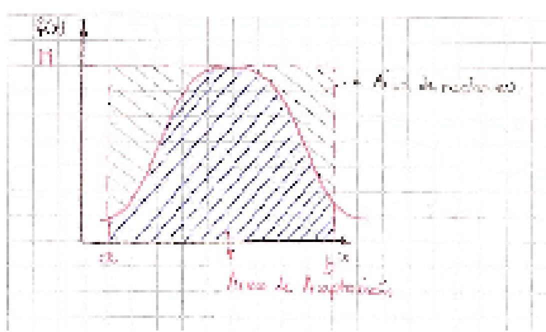
Es una función de difícil transformación, para el método anterior.
Este método parte de la generación de dos variables aleatorias r_1 y r_2, que nos permiten definir:

$x_i = a + (b - a)r_1$
$y_1 = M.r_2$ donde M es el máximo valor que puede tomar la función f(x) en el intervalo (a,b)

Si verificamos que $M.r_2 \leq f(a + (b - a).r_1)$, es decir: $y_i \leq f(x_i)$, entonces aceptamos el x_i generado, de lo contrario es rechazado.



La ventaja de este método es que no existe la necesidad de hallar la acumulada y por ende tampoco su transformada.



Es posible calcular la eficiencia del método a partir de:

$$e = \frac{Area\ bajo\ la\ curva}{Area\ del\ rectangulo}$$

MODELOS

Tipos de variables

1. **Exógenas:** aquellas externas al planteamiento del modelo, definen al mismo. Son las independientes o de entrada al modelo, y se supone que han sido predeterminadas y proporcionadas independientemente del sistema que se modela.

    Datos: información del sistema. Son los datos que debemos tomarlos de la realidad tal cual son dados y no se puede modificar. En general estos datos serán funciones de densidad de probabilidad. Ej: intervalo de arribo entre personas.
    Control: son las que nosotros imponemos al sistema. Son susceptibles de manipulación o control por quienes toman decisiones o crean políticas para el sistema. Ej: número de cajas a colocar en un supermercado.

2. **Endógenas:** propias del sistema. Son las dependientes y se generan dentro del modelo.

    Estado: describen el estado del sistema en cada instante de tiempo (foto). Ej: cantidad de stock de una mercadería.
    Resultado: nos las otorga el sistema, pero las definimos nosotros dependiendo del objetivo de la simulación. Son las variables de salida del sistema y son generadas por la interacción de las variables exógenas con las de estado. Ej: promedio de espera en un puesto de atención.

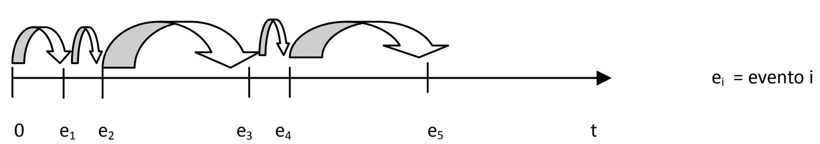
Organización temporal de procesos: constituye el mecanismo que vamos a usar para la modificación de la variable tiempo en un modelo, con el objeto de construir la historia del estado del modelo. Este mecanismo se adopta y rige durante todo el proceso.
Tenemos dos maneras de trabajarlo:

a. Evento a evento (o incrementos variables)
b. A intervalos constantes

La diferencia entre evento a evento e intervalos constantes, es que el tiempo en el primero no está explicito sino que esta tácito.

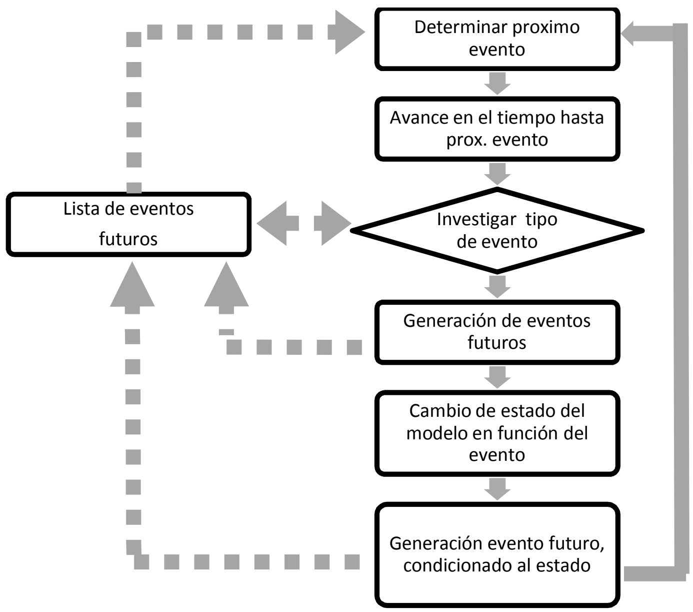
Evento a evento

Se hace pasar el tiempo centrándonos en cuándo sucede un evento. Durante el tiempo entre eventos, el estado es constante.



Si nos situamos en un evento e_i, se puede estimar el próximo evento teniendo en cuenta el estado, el evento y la historia.

Este sistema se puede utilizar cuando se tiene cierta información que me permite estimar qué pasó y cuándo pasaron los eventos.



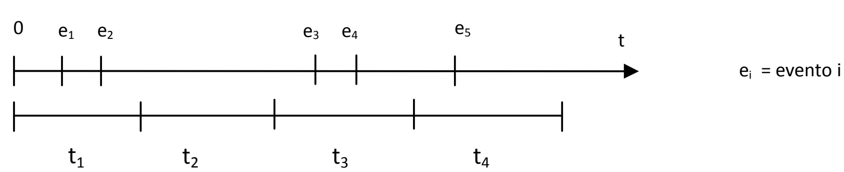
Un elemento fundamental de este procedimiento es el mantenimiento permanente de una lista de eventos futuros (considerando el tipo de evento y el tiempo en el que ocurren) de la que se obtiene el próximo evento, una vez procesado el anterior.
A dicha lista se incorpora, cuando se investiga un evento, el evento futuro generado por el evento que estamos analizando.

Evento a Evento →

- Cantidad de Personas → El tiempo de atención o servicio recién se conoce cuando la persona va a comenzar a ser atendido. Por lo tanto al llegar se coloca en la cola con menor cantidad de personas pero no necesariamente por eso saldrá antes.
- Tiempo Comprometido → El tiempo de atención o servicio se conoce cuando la persona llega. Por lo tanto al llegar se coloca en la cola donde será atendido antes, pues es posible determinar el tiempo en que se libera el servidor.

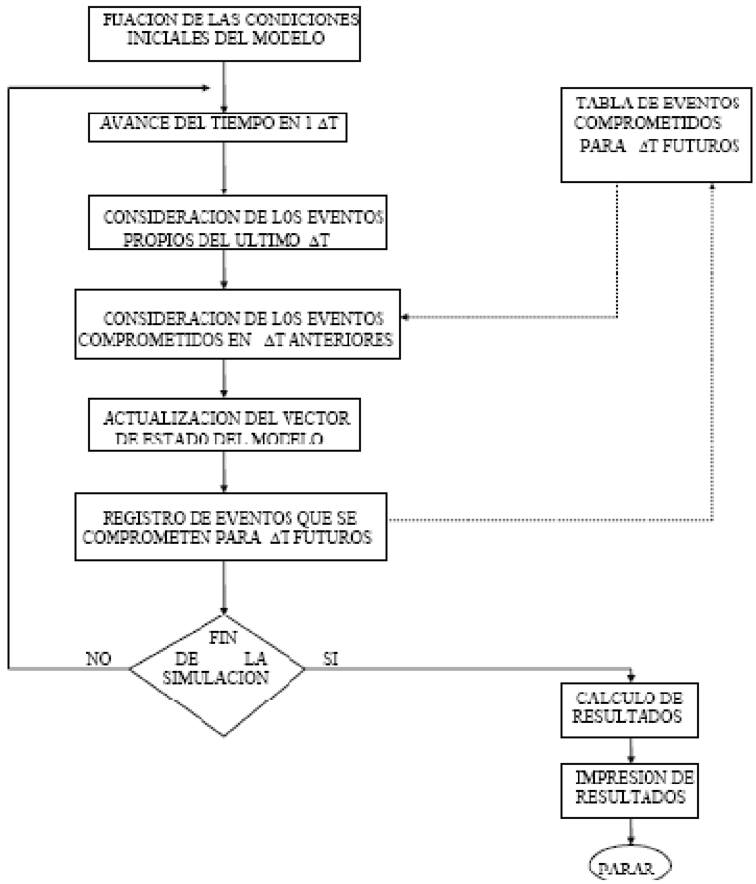
A intervalos constantes

Aquí se hace correr el tiempo, cada ciertos intervalos t constantes (o Δt), y se observa que sucede en esos puntos.



Puede suceder que en el intervalo de tiempo que observamos sucedan varios eventos, pero así también puede que no suceda ninguno.

Este modelo es utilizado en aquellos casos que se conocen de antemano cuando suceden los eventos, por ejemplo todos los días a las 2:00, evitando así el análisis que se da con el modelo evento a evento.
Se trata de fijar un Δt, t suficientemente pequeño, de tal manera que no se escape ningún evento, pero muchas veces voy a analizar si pasó algo y no sucedió nada.



Modelo de inercia

Las variables de control toman un stock medio o de referencia (SR) del cual partirá el análisis.
Se llama inercia porque permite ir actualizando el tamaño del pedido dinámicamente.

$$TP = SMax - \frac{SA_1 + SA_2 + SA_3 + \cdots + SA_n}{n} \qquad SA_1 = ST - VA - VP$$

Donde:

- TP: tamaño de pedido
- VA: ventas atrasadas
- VP: ventas perdidas
- SMax: stock de máxima reposición

$SA_1 + SA_2 + SA_3 + \cdots + SA_n$ → indica cuantos periodos anteriores se consideran, en este caso n

Cuanto mayor sean los periodos considerados, mayor será la inercia.

Si se promedian muchos periodos anteriores, las grandes o bruscas variaciones no se reflejan. Dicho de otra manera, los cambios bruscos no influyen demasiado cuando se toman muchas SA_i, es decir cuando se tiene bastante historial (periodos de tiempo).
Si el sistema es inherentemente inestable, y se toman muchos periodos en cuenta, entonces no se reflejan rápidamente los cambios, lo cual es malo. En consecuencia, si se tienen picos constantes en nuestro modelo, es recomendable tener una inercia menor.

Análisis: $SA_1 = ST - VA - VP$

- Si ST ≠ 0 → VA = VP = 0 → esto va reduciendo el TP
- Si ST = 0 → VA ≠ VP ≠ 0 → esto va aumentando el TP

Simulaciones terminales

Los modelos de tipo terminal tienen como característica principal la ocurrencia de un evento que da por terminada la simulación.

Intervalos de confianza

Debido a la naturaleza aleatoria de los resultados de este tipo de modelos, es necesario determinar su distribución de probabilidad y su intervalo de confianza en las diferentes replicas.
Si la variable aleatoria sigue una distribución normal, el intervalo de confianza está dado por:

$$IC_{dist.normal} = \left[\bar{x} - \frac{s}{\sqrt{r}}\left(t_{\frac{\propto}{2},r-1}\right); \bar{x} + \frac{s}{\sqrt{r}}\left(t_{\frac{\propto}{2},r-1}\right)\right]$$

Para otro tipo de distribución:

$$IC_{otra\ dist.} = \left[\bar{x} - \frac{s}{\sqrt{r * \propto/2}}; \bar{x} + \frac{s}{\sqrt{r * \propto/2}}\right]$$

Donde:
- r: número de réplicas
- α: nivel de rechazo

$$\bar{x} = \frac{1}{r}\sum_{i=1}^{r} x_i$$

$$s = \sqrt{\frac{1}{r-1}\sum_{i=1}^{r}(x_i - \bar{x})^2}$$

Simulaciones no terminales o de estado estable

No involucran una ocurrencia en el tiempo en que tengan que finalizar.

Longitud de las réplicas

Para que el resultado de una variable aleatoria llegue al estado estable en una simulación no terminal, es necesario garantizar que la longitud de la réplica (n) sea lo suficientemente grande para que la variación entre réplicas no difiera de cierta exactitud (Є), el 100(1-α)% de las veces.
En caso de normalidad el tamaño de la corrida se calcula como:

$$n = \left(\frac{\sigma . Z_{\alpha/2}}{\epsilon}\right)^2$$

σ: desvío estándar
Є: rango de variación

Si se tiene la certeza de normalidad pero se desconoce el valor de la desviación estándar, será necesario realizar una corrida inicial de tamaño n’ para determinar un estimador de la desviación. En este caso la longitud de la réplica se determina mediante:

$$n = \left(\frac{S}{\epsilon}\left(t_{\frac{\alpha}{2}, n'-1}\right)\right)^2$$

DINAMICA DE SISTEMAS

El objetivo básico de la dinámica de sistemas es llegar a comprender las causas estructurales que provocan el comportamiento del sistema.
Como característica diferenciadora de otras metodologías puede decirse que no se pretende predecir detalladamente el comportamiento futuro.
Otra característica importante es su enfoque a largo plazo.
En este caso el ajuste del modelo a los datos históricos ocupa un lugar secundario, siendo el análisis de la lógica interna y de las relaciones estructurales en el modelo los puntos fundamentales de la construcción del mismo.
Se analiza el comportamiento de las distintas partes del sistema a lo largo del tiempo. Permite conocer sus partes y como se relacionan (sus vínculos).
Se tiene:

- **Análisis del sistema** = reconocer el comportamiento de las partes
- **Síntesis del sistema** = estudio de cómo se relacionan las partes

El análisis de un sistema, consiste en su concepción en forma individual para conocer las partes que lo forman. Sin embargo, el mero análisis del sistema no es suficiente, no basta saber cuáles son sus partes, de manera tal de comprender su comportamiento necesitamos saber cómo se integran, cuales son los mecanismos mediante los que se produce su coordinación, necesitamos entonces saber cómo se produce las síntesis de las partes del sistema.
En dinámica de sistemas, vamos a ocuparnos de analizar como las relaciones, en el seno de un sistema, permiten explicar su comportamiento. Un sistema es un conjunto de elementos en interacción, y esta interacción es el resultado de que algunas partes influyan sobre otras. Estas influencias mutuas determinarán cambios en esas partes, por lo tanto los cambios que se producen en el sistema son reflejo en alguna medida de las interacciones que se tienen en su seno. La trama de relaciones constituye lo que se denomina **estructura**.

¿Para qué se aplica la dinámica de sistemas?

Para el análisis de situaciones donde los sistemas se ven influenciados por casos de retroalimentación, o casos de comportamientos aleatorios.
El uso de esta metodología es muy amplio, desde la zoología hasta la toma de decisiones gerencial.

Existen dos modelos de hechos futuros:

- **Modelo de predicción:** busca otorgar un valor preciso para tomar la decisión.
- **Modelo de gestión:** lo único que hace es asegurar, entre alternativas, cual es mejor o peor que la otra, pero no necesariamente la óptima.

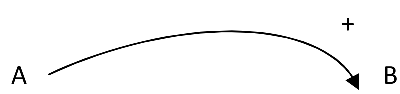
Pasos:

1. **Determinar cuál es el problema:** se busca identificar el problema con claridad y describir los objetivos con precisión. ¿Qué costo nos insume? ¿Qué tiempo? ¿Los beneficios justifican la investigación? ¿Es suficiente este modelo de simulación?

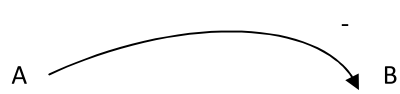
2. **Definir el sistema:** cuál es el sistema sobre el cual vamos a trabajar, sabiendo que un sistema es un conjunto de elementos relacionados entre sí, de forma tal que un cambio en uno de los elementos afecta a todos. ¿Cuáles son los elementos relacionados directa o indirectamente al sistema?
3. **Límite del sistema:** el sistema debe contener el menor número de elementos posibles, que nos permitan realizar una simulación para explicar al final, cuál de las propuestas de aplicación que hemos planteado y estudiado es la más eficaz para resolver el problema. En la construcción del modelo, se añaden y suprimen elementos en función a este concepto.

Diagrama de causal

El conjunto de elementos que tienen relación con nuestro problema y permiten en principio explicar el comportamiento observado, junto con las elaciones entre ellos (en muchos casos de retroalimentación), forman el sistema. El **diagrama de causal** es un diagrama que recoge los elementos clave del sistema y las relaciones entre ellos.
Las diferentes relaciones están representadas por flechas entre las variables afectadas por ellas.
Estas flechas van acompañadas de un signo (+ o -) que indica el tipo de influencia ejercida por una variable sobre la otra. Un signo '+' quiere decir que un cambio en la variable origen de la flecha producirá un cambio en el mismo sentido en la variable destino. El signo '-' simboliza que el efecto producido será en sentido contrario.
Cuando un incremento de A, produce un incremento de B, o bien una disminución de A provoca una disminución de B, se tendrá una **relación positiva**.



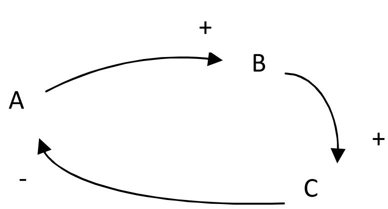
Cuando un incremento de A, produce una disminución de B, o bien una disminución de A provoca un aumento de B, se tendrá una **relación negativa**.

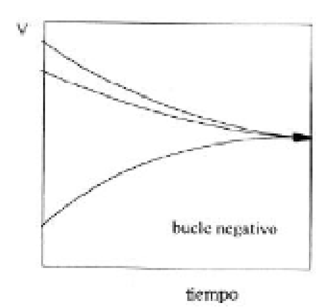


Retroalimentación

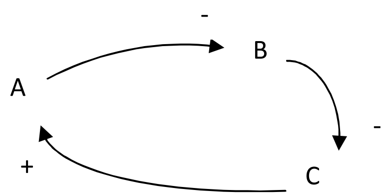
Bucle: es una cadena cerrada de relaciones causales.

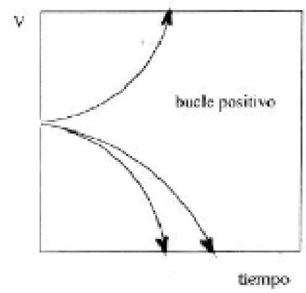
Bucles negativos: actúan como estabilizadores de los sistemas, al dirigirlos hacia un objetivo determinado. El número de relaciones negativas es impar.





Bucles positivos: la cantidad de relaciones negativas es par. Tienden a hacer inestable el sistema. Causan crecimiento, evolución y también el colapso de los sistemas.





Factor limitativo: es aquel elemento del sistema que **ahora** limita el crecimiento del sistema, o en un futuro inmediato. Es único en cada momento, pero a lo largo del tiempo distintos elementos del sistema pueden actuar como factores limitativos. El factor limitativo es dinámico. Ej: cantidad de materia prima, capacidad de mano de obra.

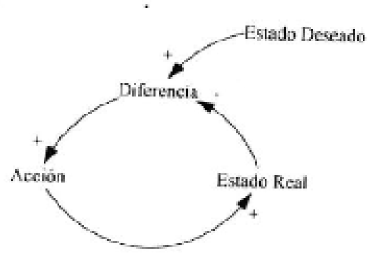
Factores claves: son aquellos que podemos utilizar para generar o conseguir grandes cambios en el sistema con un esfuerzo mínimo. Los factores clave (key factors or leverage points) son elementos del sistema a los que este es muy sensible. Siempre son los mismos, no suelen variar a lo largo del tiempo. Provocan reacciones violentas. Es importante conocerlos si deseamos manipular el estado del sistema, evitando alterar aquellos que provocarán una reacción negativa del sistema, y en cambio trataremos de aprovechar aquellos que van a provocar una reacción favorable. Es importante recordar que en general se hallan ocultos y que son siempre los mismos.

Tipos de sistemas

La estructura interna determina el comportamiento de los sistemas.

Sistemas estables e inestables

Un sistema es **estable** cuando se halla formado o dominado por un bucle negativo, y es **inestable** cuando el bucle es positivo. Es decir, cuando en el bucle dominante haya un número impar de relaciones negativas, tendremos un bucle negativo, y el sistema será estable. La estructura básica de los sistemas estables está formada por un **Estado Deseado** y por un **Estado Real** del sistema, estos dos estados se comparan (**Diferencia**), y en base a este valor el sistema toma una **Acción** para igualar el estado real al deseado.
El sistema compara permanentemente su estado real con el deseado, y cuando existe una diferencia, hace acciones en el sentido de acercar su estado real al deseado.



Sistemas hiperestables

Cuando un sistema está formado por múltiples bucles negativos, cualquier acción que intenta modificar un elemento no se ve contrarrestado solo por el bucle en el que se halla dicho elemento, sino por todo el conjunto de bucles negativos que actúan en su apoyo, super-estabilizando el sistema.

Sistemas oscilantes

Para que un sistema muestre un comportamiento oscilante es necesario que tenga al menos dos “niveles”, que son elementos del sistema en los que se producen acumulaciones.
Si el estado actual del sistema no nos gusta o no es el correcto, no es necesario hacer nada ya que todo parece ser cíclico y volverá a la normalidad por sí solo.

Sistemas sigmoidales

Son sistemas en los cuales existe un bucle positivo que actúa en un principio como dominante y hace arrancar el sistema exponencialmente, y después el control del sistema lo toma un bucle negativo que anula los efectos del anterior y proporciona estabilidad al sistema, situándolo en un valor asintóticamente.

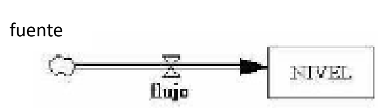
<u>Diagrama de flujo o de Forrester</u>

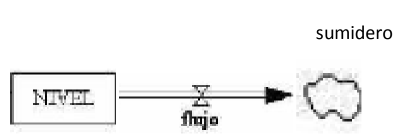
Es una traducción del diagrama causal a una terminología que facilita la escritura de las ecuaciones en el ordenador.
No hay reglas precisas de cómo hacer la transformación, pero se pueden seguir los siguientes pasos:

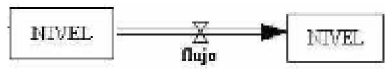
1. Realizar una fotografía mental al sistema y lo que salga de él, esos son los **niveles**.
2. Buscar o crear unos elementos que sean “la variación de los niveles”, esos son los **flujos**.
3. El resto de los elementos son **variables auxiliares**.

Niveles: son aquellos elementos que nos muestran en cada instante la situación del modelo, presentan una acumulación y varían sólo en función de otros elementos llamados flujos. Las “nubes” dentro del diagrama de flujo son niveles de contenido inagotable. Los niveles se representan con un rectángulo.

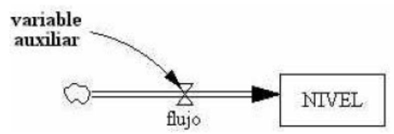
Flujos: son elementos que pueden definirse como funciones temporales. Puede decirse que recogen las acciones resultantes de las decisiones tomadas en el sistema, determinando las variaciones de los niveles.





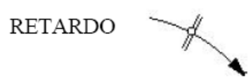


Variables auxiliares: y las constantes, que son parámetros, permiten a una visualización mejor de los aspectos que condicionan el comportamiento de los flujos.



Las magnitudes físicas entre flujos y niveles se transmiten a través de los denominados “canales materiales”. Por otra parte, existen los llamados “canales de información” que transmiten información que por su naturaleza no se conservan.

Por último los “retardos” simulan los retrasos de tiempo en la transmisión de los materiales o las informaciones. En cierta forma los retardos de información actúan como filtros alisadores de la variable de entrada.



Los modelos de dinámica de sistemas no son modelos predictivos, no pretenden hallar valores exactos, sino comparativos, es decir han de permitir comparar diferentes políticas alternativas en base al escenario al que conducen.