

INTRODUCCIÓN

Es muy importante la modelización, pero la simulación no es sólo eso. Hay que derivar conclusiones válidas estadísticamente de las repeticiones que se realizan de los modelos programados. El error principal de los estudios de simulación ha sido la incorrecta experimentación con dichos modelos, lo que ha llevado a la incorrecta estimación de los “verdaderos” parámetros de las distribuciones subyacentes a los sistemas simulados.

Si: n =número de repetición,

m =número de muestra dentro de una repetición,

entonces, los resultados de un experimento podrían ser:

$$Y_{11}, \dots, Y_{1i}, \dots, Y_{1m}$$

$$Y_{21}, \dots, Y_{2i}, \dots, Y_{2m}$$

...

$$Y_{n1}, \dots, Y_{ni}, \dots, Y_{nm}$$

¿Cómo estimar las variables aleatorias Y_m subyacentes?

Los problemas asociados con la estimación de las variables son los siguientes:

1. Independencia
2. Sesgo
3. Régimen permanente vs. transitorio

TIPOS DE VARIABLES A ESTIMAR

Basadas en Observaciones.

El valor que se recoge de la simulación es un valor puntual, esto es, recoge el valor de la variable en un momento específico de la simulación.

Ejemplo: Tiempo de estancia en la cola

El valor se estima asumiendo normalidad mediante la media y la desviación típica:

$$\bar{W} = \frac{\sum_{i=1}^n W_i}{n}$$

$$\sigma_w = \sqrt{\frac{\sum_{i=1}^n (W_i - \bar{W})^2}{n}}$$

Basadas en el Tiempo.

El valor se calcula como media ponderada a lo largo de todo el período de estimación

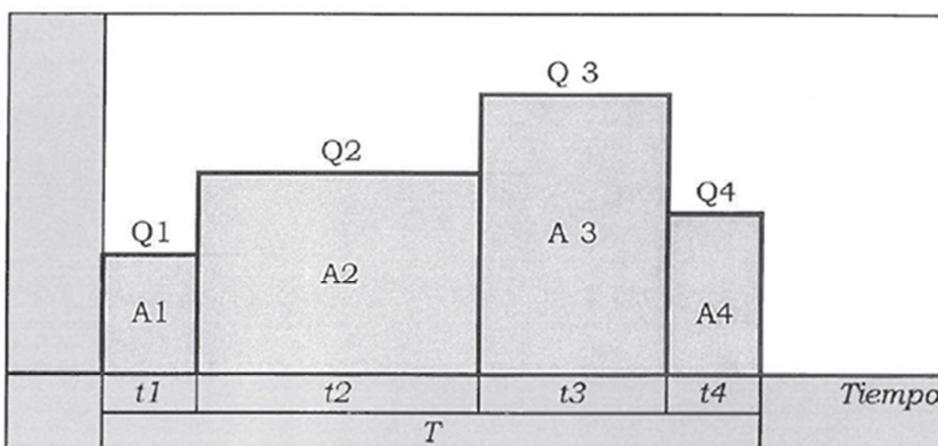
Ejemplo: Longitud de la cola

$$\bar{Q} = \frac{\sum_{i=1}^n Q_i t_i}{n}$$

$$\sigma_Q = \sqrt{\frac{\sum_{i=1}^n (Q_i - \bar{Q})^2 t_i}{T}}$$

EJEMPLO

Calcular la longitud media y máxima de cola, si lo sucedido hasta el tiempo T de simulación es lo siguiente:



a) Basada en observación Longitud Máxima de Cola = Q3

b) Basada en el tiempo:

$$\text{Longitud Media de Cola} = \bar{Q} = \frac{\sum_{i=1}^n Q_i t_i}{n} = \frac{A_1 + A_2 + A_3 + A_4}{n}$$

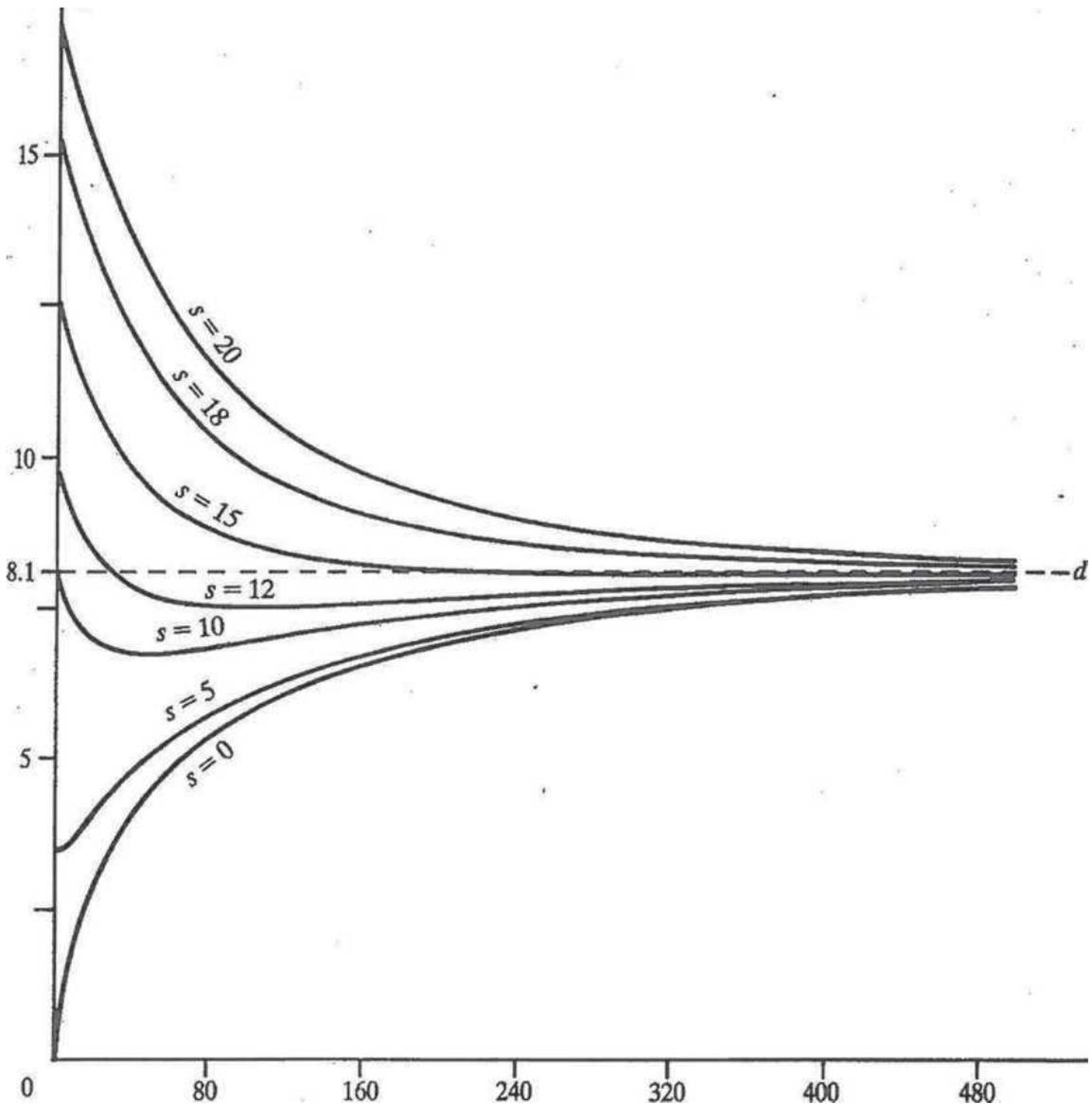
RÉGIMEN PERMANENTE

- Es el estado "normal" de funcionamiento del sistema, cuando el sistema está bajo control y el comportamiento de las variables del sistema se puede conocer.

- No significa que todas las muestras que se tomen de una variable tengan el mismo valor, sino que sigan la misma distribución.

- Dicha distribución no depende de las condiciones iniciales del sistema. El ritmo de convergencia hacia las condiciones iniciales, en cambio, sí que depende de dichas condiciones iniciales.

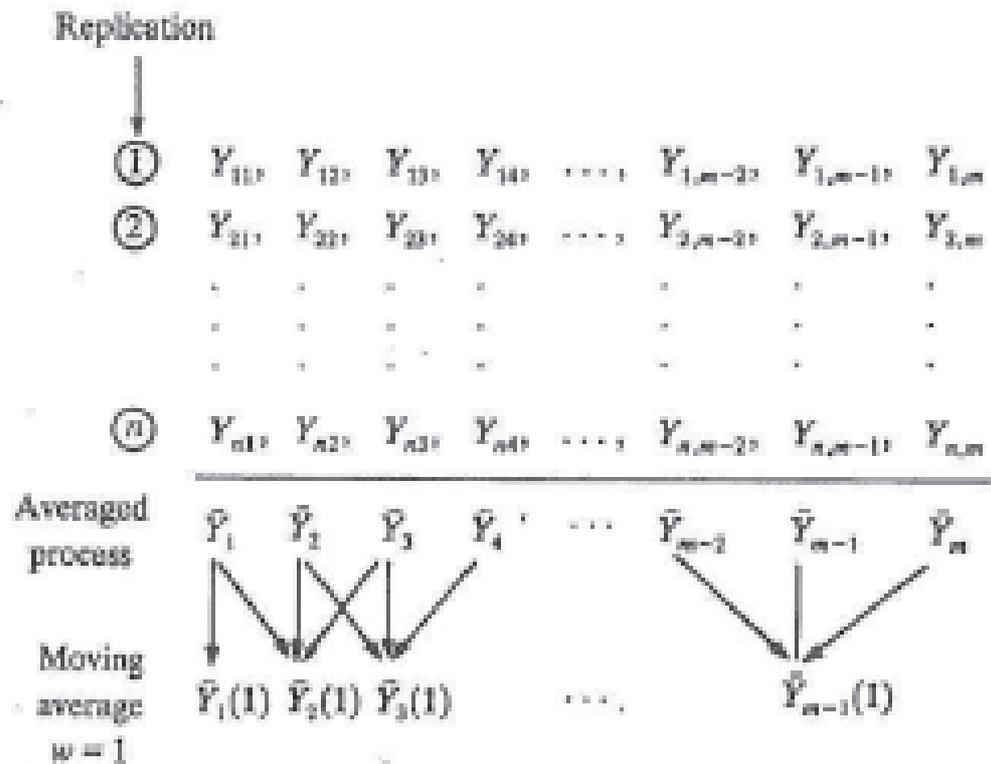
$E(D)$



- El período de calentamiento de un modelo sirve para eliminar aquellas observaciones que se tomen sin haberse conseguido las condiciones de régimen permanente. Estas observaciones "prematuras" provocan sesgo debido a la influencia de las condiciones iniciales en la estimación de los parámetros de régimen permanente.

• Procedimiento de Welch para determinar la longitud del período de calentamiento (**lpc**):

1. Se realizan **n** repeticiones de **m** observaciones



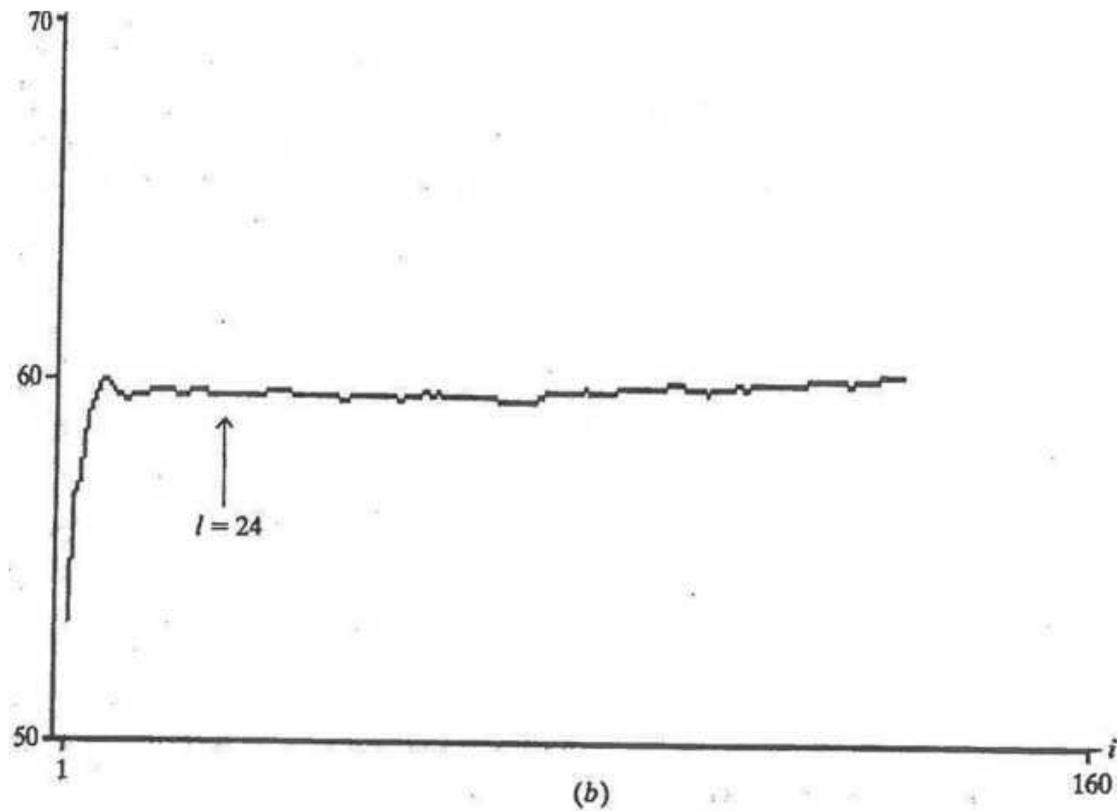
2. Se halla la media de cada observación

3. Se calcula la media móvil de las medias, donde **w** = ventana, o número de muestras anteriores y posteriores que se incluyen en la media móvil

$$\bar{Y}_i(w) = \begin{cases} \frac{\sum_{s=i-w}^i \bar{Y}_{t+s}}{2w+1} \\ \forall i = w+1, \dots, m-w \end{cases}$$

$$\bar{Y}_i(w) = \begin{cases} \frac{\sum_{s=0}^{i-1} \bar{Y}_{t+s}}{2i+1} \\ \forall i = 1, \dots, w \end{cases}$$

4. Se halla gráficamente la longitud del período de calentamiento, l_{pc} :



TIPOS DE SIMULACIÓN CON RESPECTO AL ANÁLISIS DE RESULTADOS

DEPENDIENDO DE LA LONGITUD DE UNA REPETICIÓN

Con final: existe un evento “natural” que indica el final de la repetición, y tras el cual se produce una limpieza del sistema. Es fundamental la estimación de las condiciones iniciales pues influyen directamente en el análisis del sistema.

Ej: Apertura y cierre de establecimientos (de 8 a 3)

Ej: Producción de N unidades.

Sin final: no existe el mencionado evento “natural”. Las operaciones de un período de tiempo influyen en el siguiente, pues las condiciones iniciales de una repetición son las condiciones finales del período anterior.

Ej: Una cadena de producción.

ANÁLISIS ESTADÍSTICO DE SIMULACIONES CON FINAL

- Estimación de la media y de la varianza

$$\bar{X}(n) = \frac{\sum_{i=1}^n X_i}{n} \quad S^2(n) = \frac{\sum_{i=1}^n (X_i - \bar{X}(n))^2}{n-1}$$

- Intervalo de confianza “aproximado” de la media

$$\bar{X}(n) \pm t_{n-1, 1-\alpha/2} \sqrt{\frac{S^2(n)}{n}}$$

- Relación “número de repeticiones-precisión”

$$\delta(n) = t_{n-1, 1-\alpha/2} \sqrt{\frac{S^2(n)}{n}}$$

- Estimación de percentiles

$$\hat{x}_q = \begin{cases} X_{(nq)} & \\ \text{si } nq \text{ es entero} & \\ X_{(\lfloor nq+1 \rfloor)} & \end{cases}$$

Estimación de condiciones iniciales:

1. Empezar siempre con un sistema vacío (Banco, a las 08:00)
2. Recoger datos en el tiempo de inicio de la simulación y empezar con el sistema ya “en funcionamiento”

Resumen: N repeticiones con suma importancia de las condiciones iniciales

ANÁLISIS ESTADÍSTICO DE SIMULACIONES SIN FINAL

- 6 métodos distintos de estimación de la media

1. Método de Repetición/Eliminación
2. Método de las Medias por Lotes
3. Método Autorregresivo
4. Método Espectral
5. Método Regenerativo

6. Método de las Series Temporales Estandarizadas

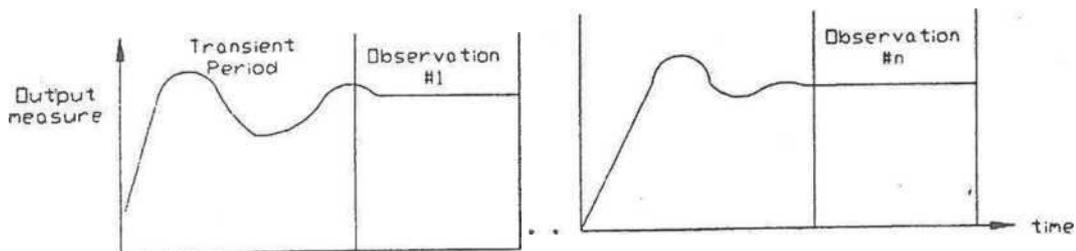
- Todos los métodos buscan muestras diferentes de un mismo modelo para poder estimar estadísticamente de forma correcta las características principales del sistema representado.
- Una vez conseguidas las muestras, las estimaciones de la media se realizan de la siguiente forma:

1. Media

2. Intervalo de Confianza

Método de Repetición/Eliminación

- Metodología:

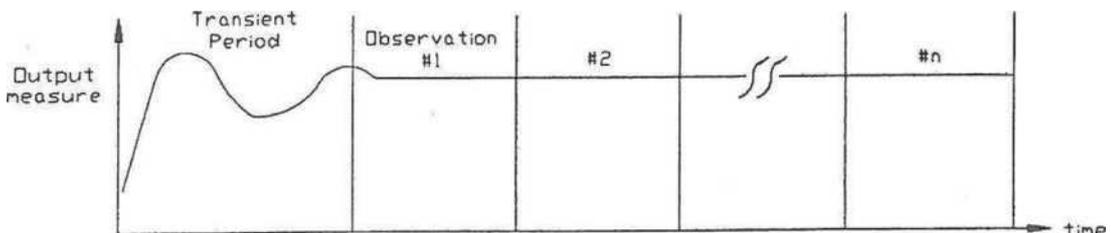


- Ventajas:

1. Produce buenos resultados estadísticamente
2. Es el método más fácil de aplicar y de entender
3. Se puede aplicar a cualquier tipo de parámetro
4. Se puede utilizar para estimar varios parámetros del mismo modelo
5. Se puede utilizar para comparar escenarios diferentes
6. Problema: potencial sesgo en la estimación de la media
7. Dificultad: estimación del período de calentamiento

Método de las Medias por Lotes

- Metodología:



- Ventaja: sólo se ejecuta una vez el período de calentamiento
- Problema: potencial sesgo en la estimación de la varianza

- Dificultad: estimación del período de lote, para que no exista correlación entre las muestras.

Resumen: 1 repetición, con suma importancia en la estimación del período de lote