

# **Universidad Politécnica de Cartagena**



## **Escuela Técnica Superior de Ingeniería de Telecomunicación**

### **PRÁCTICAS DE MODELADO Y SIMULACIÓN**

#### **Práctica 4: Análisis estadístico de resultados**

Profesores:

Javier Vales Alonso

## 1. INTRODUCCIÓN

En esta práctica se implementará en matlab un **sistema de análisis estadístico** de los resultados obtenidos con el simulador. Este sistema se integrará en el esqueleto de simulación y se encargará de finalizar las simulaciones cuando hayan alcanzado un nivel de calidad suficiente.

## 2. OBJETIVOS

Relativos a simulación:

- Implementar un sistema de análisis estadístico de los resultados. Este sistema filtrará las muestras durante el régimen transitorio, efectuará un filtrado por bloques de las mismas para evitar correlaciones a corto plazo en las muestras y determinará el intervalo de confianza para el resultado.
- Crear un sistema que permita finalizar las simulaciones cuando éstas alcancen un nivel de calidad adecuado

Relativos al sistema bajo estudio:

- Observar las dependencias entre los parámetros de configuración del simulador y los resultados de naturaleza estadística alcanzados.

## 3. DESARROLLO DE LA PRÁCTICA

La práctica se divide en cuatro partes:

1. Primero se eliminarán las muestras transitorias.
2. A continuación se realizará un muestreo por bloques.
3. Después se calculará la calidad asociada al intervalo de confianza generado para las muestras obtenidas.
4. Finalmente, se integrará el proceso de obtención de calidad en las condiciones de parada del simulador y lo probará en distintas configuraciones prácticas, con el fin de obtener resultados sobre el comportamiento del sistema.

**La metodología de análisis que implementaremos es independiente del sistema simulado.**

A modo de ejemplo, trabajaremos en esta práctica con la estructura de simulación para colas G/G/k creada en la práctica anterior. Posteriormente, deberá aplicar esta misma metodología para el trabajo de simulación propuesto.

Los procedimientos aquí empleados se describen en detalle en el Tema 6 de la asignatura, es importante su estudio antes de proceder a la realización de la práctica.

### 3.1. Filtrado de muestras transitorias

Las muestras iniciales de un simulador deben ser descartadas para evitar sesgos en el resultado. **La idea es simplemente descartar las primeras H muestras de la simulación.** Defina ese parámetro en la cabecera de su simulador:

```
H = 100000; % Muestras transitorias, se descartan
```

### 3.2. Implementación del muestreo por bloques

Como estudiamos en teoría, habitualmente el proceso simulado experimenta autocorrelación a corto plazo y es necesario “corregir” las muestras. Para ello desarrollaremos el método estudiado de análisis por bloques. Éste es muy sencillo, y **consiste en coger una muestra y descartar las (D-1) muestras siguientes**.

`D = 10; % Tamaño del bloque, se coge sólo una de cada D muestras`

### 3.3. Implementación de la función *calidad*

Nuestro siguiente objetivo es calcular, dado un conjunto de muestras, el intervalo de confianza para el parámetro buscado y determinar la probabilidad (la *calidad* o nivel de confianza) de que el verdadero valor del parámetro (que es desconocido) se encuentre dentro de ese intervalo. Usaremos la siguiente notación:

1. Muestras (son **variables aleatorias idénticamente distribuidas**):

$$\{X_1, \dots, X_n\}$$

2. Media *poblacional* (parámetro a estimar, es determinista):

$$\mu_X = E\{X_i\}$$

3. Varianza *poblacional* (determinista):

$$\sigma_X^2 = E\{(X_i - \mu_X)^2\}$$

Nótese que las dos expresiones anteriores son consistentes con la hipótesis señalada de que las N muestras son idénticamente distribuidas, y por tanto tendrán misma media y varianza.

4. Media *muestral* (estimador mínimo-cuadrático de la media, es una **variable aleatoria**):

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

5. Cuasi-Varianza *muestral* (estimador *insesgado* de la varianza, es una **variable aleatoria**):

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

6. **Calidad** (1- $\alpha$ ) es la probabilidad de que la media poblacional se encuentre dentro de un intervalo de confianza centrado en la media muestral y cuya semi-longitud se denomina **tolerancia** (t). Es decir:

$$p(\bar{X}_n - t < \mu_X \leq \bar{X}_n + t) = 1 - \alpha$$

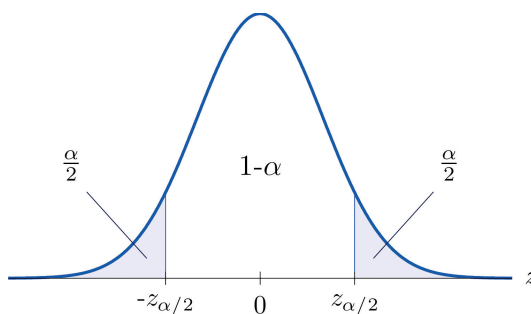
La tolerancia se debe expresar en % para evitar distorsiones debido a la magnitud de la media muestral. En este caso se conoce como tolerancia relativa ( $t\%$ ) y la absoluta se calcula como:

$$t = t\% \bar{X}_n$$

**7. Tolerancia normalizada.** Se denota  $z_{\alpha/2}$  y su relación con la tolerancia está dada por:

$$z_{\alpha/2} = \frac{t}{\sqrt{\frac{S_{n-1}^2}{n}}}$$

Visualmente, puede verse la **relación entre la tolerancia normalizada y la calidad** en la distribución Normal (0,1):



Como se ha indicado en teoría, a partir de la tabla de la distribución Normal (0,1), y dada  $z_{\alpha/2}$  se obtiene el valor de  $1-\alpha$  correspondiente y viceversa.

**EJERCICIO 1:** Expresar en función de  $z_{\alpha}$  la ecuación para la calidad dada en el punto 6.

**EJERCICIO 2:** Durante la simulación no es posible (ni necesario) guardar todas las muestras, basta con disponer de las siguientes variables acumuladas:

nummuestras  $\rightarrow n$

summuestras  $\rightarrow \sum_{i=1}^n X_i$

sumcuadrado  $\rightarrow \sum_{i=1}^n (X_i)^2$

Expresar en función de estas 3 variables (`nummuestras`, `summuestras`, `sumcuadrado`) la fórmula de la cuasi-varianza muestral del punto 5

**EJERCICIO 3:** Debe crear una función *calidad* que, dadas las 3 variables anteriores, calcule la calidad asociada al intervalo de confianza de tolerancia relativa dada

```
function [unomenosalfa, intizqda, intderecha] =
    calidad(tolrelativa, nummuestras, summuestras, sumcuadrado)
% Calcula el intervalo de confianza asociado a las muestras y
% Devuelve la probabilidad asociada al intervalo de confianza

end
```

En matlab podemos obtener fácilmente el valor del  $\alpha$  asociado al valor de  $z_{\alpha/2}$ . El método es:

```
normcdf(z) % Devuelve la probabilidad acumulada hasta el punto z
```

Por tanto:

```
1 - normcdf(z) % Probabilidad en la cola a la derecha de z
```

Como estamos buscando la probabilidad de que  $z$  quede más a la derecha de  $z_{\alpha/2}$  o a la izquierda de  $-z_{\alpha/2}$  y la normal es simétrica será:

```
alfa = (1 - normcdf(z))*2
```

Finalmente,

```
unomenosalfa = 1 - (1 - normcdf(z))*2
```

### 3.4. Parada del simulador atendiendo a criterios de calidad

Por último, vamos a emplear el método de obtención de calidad para fijar las condiciones de parada del simulador. En vez de dar un número fijo de pasos, se parará la simulación cuando ésta alcance una calidad mínima. El procedimiento a implementar es el **Algoritmo 3 del Tema 6, “Simulación de procesos ergódicos CCP”**, debe estudiarlo en detalle antes de proceder con su implementación.

Debe añadir en su simulador los parámetros:

```
calidadobjetivo = 0.99; % Suele estar entre el 90% y el 99%
tolrelativa = 0.05; % Suele estar entre el 1% y el 10%
TEST = 1000; % Cada cuantas muestras comprobamos calidad
```

A continuación, sustituya el bucle **for** de simulación por la siguiente estructura:

```
pasos = 0;
while true
    pasos = pasos+1;

    % EL INTERIOR DEL BUCLE NO CAMBIA

    % ELIMINAMOS TRANSITORIO Y FILTRAMOS POR BLOQUE.
    % CUANDO RECOJAMOS "TEST" NUEVAS MUESTRAS COMPROBAMOS LA CALIDAD:

    if ~mod(nummuestras,TEST)
        [unomenosalfa, tizq, tdrcha] = calidad(...);
        if (unomenosalfa >= calidadobjetivo) break; % FIN SIMULACION
    end

    % SI MEDIMOS VARIOS PROCESOS DEBEMOS OBTENER LA CALIDAD DE CADA UNO Y EL
    % IF DEBE COMPROBAR SI LA CALIDAD ES MAYOR A LA DESEADA PARA CADA PROCESO
end
```