**CAPÍTULO 3**

**3. ESTIMACIÓN JACKNIFE UTILIZANDO SIMULACIÓN**

**3.1 Introducción**

En este capítulo desarrollamos el modelo de simulación del sistema identificado y planteado, además se explican cada una de las funciones desarrolladas en el programa de simulación, el cual se realizó en el lenguaje Matlab 5.3; así tenemos en la sección 3.2 el desarrollo del modelo de simulación, en la sección 3.3 se encuentran la descripción de los subprogramas realizados en el simulador, tales como, la subfunción Variable en la sección 3.3.1, en la sección 3.3.2 la subfunción Resultados, en la sección 3.3.3 la subfunción Procedimiento\_Operaciones, en la sección 3.3.4 la subfunción Generar\_Muestra, en la sección 3.3.5 la subfunción Estimador\_Jacknife y en la sección 3.3.6 la subfunción Gráficos.

**3.2 Modelo de simulación**

El sistema a simular es dinámico ya que sus componentes tienen un comportamiento aleatorio, el objetivo es explorar dicho sistema, comparando los dos métodos de estimación: Jacknife y Convencional.

El proceso consiste en que el investigador seleccione dependiendo de su necesidad, una población con sus respectivos parámetros poblacionales y un estimador de algún parámetro de la misma, se generan 50 muestras aleatorias de tamaño n de la población estipulada y a partir de éstas obtenemos: el valor del estimador para cada una de las muestras, su sesgo e intervalo de confianza al 95%, tanto por el método Jacknife como por los métodos convencionales o usualmente utilizados de estimación.

Con los 50 estimadores y para cada uno de los métodos, calculamos la media del estimador, varianza, sesgo de simetría, mínimo, máximo, y sesgo de estimación promedio. Además graficamos el histograma de frecuencia relativa de los estimadores obtenidos por cada uno de los métodos. El tamaño muestral, n con que se trabajará es de 5, 10, 50, 100, y 500.

Con los resultados de esta simulación podremos observar la bondad de éstos métodos de estimación para los parámetros poblacionales en las situaciones dadas.

Para establecer el modelo de simulación debemos identificar sus elementos, teniendo así:

Actividad: Obtención de las medidas como el sesgo, media y varianza de los estimadores obtenidos a través de los métodos Jacknife y convencional.

Entidades: Identificamos como entidad a la unidad observada de la población teórica.

Atributos: Los atributos para la entidad unidad observada de la población son:

* Valor de la unidad observada: uk,i
* Tamaño de la muestra: n
* Número de “vueltas” (# de estimadores a generarse a partir de las muestras): nv
* Parámetro Poblacional: parametro
* Vector con la muestra aleatoria i-ésima: mi
* Valor del estimador i-ésimo Jacknife: VEJi
* Valor del i-ésimo estimador obtenido por el método convencional: VECi
* Intervalo de confianza para el valor del estimador Jacknife i-ésimo: ICJi
* Intervalo de confianza para el valor del estimador i-ésimo obtenido por el método convencional: ICCi
* Sesgo obtenido con el valor del estimador Jacknife i-ésimo: SJi
* Sesgo obtenido con el valor del estimador i-ésimo obtenido por el método convencional: SCi
* Media del estimador obtenido con el método Jacknife: MJ
* Media del estimador obtenido con el método convencional: MC
* Sesgo promedio obtenido con el método Jacknife: SPJ
* Sesgo promedio obtenido con el método convencional: SPC
* Varianza del estimador obtenido con el método Jacknife: VEJ
* Varianza del estimador obtenido con el método convencional: VEC

Teniendo así las siguientes relaciones funcionales.





El programa para el modelo está desarrollado bajo la plataforma de Matlab 5.3, y se ha utilizado programación modular para una ejecución más flexible potente y rápida del programa, la subdivisión es la siguiente: la función principal, Prog\_principal donde el investigador puede elegir cualesquiera de las poblaciones y estimadores a simular.

Además se desarrollaron subfunciones que son utilizadas en el simulador, las cuales se explican, en la siguiente sección.

**3.3 Subprogramas realizados en el simulador**

 **3.3.1 Subfunción Variables**

 Está función es llamada por la función Prog\_principal, y crea las variables necesarias para el ingreso de los parámetros poblacionales, que serán utilizados en el programa.

Las poblaciones que puede seleccionar el investigador son:

**Cuadro 1**

*Estimación por el Método Jacknife*

**Poblaciones Discretas y Continuas utilizadas en la simulación**

|  |  |
| --- | --- |
| **Discretas** | **Continuas** |
| * Binomial
* Binomial Negativa
* Poisson
* Hipergeométrica
 | * Uniforme
* Exponencial
* Beta
* Normal
* Normal Bivariada
 |

**Elaboración**: *R. Plúa*

Los estimadores que puede seleccionar el investigador para las poblaciones indicadas son los siguientes:

**Cuadro 2**

*Estimación por el Método Jacknife*

**Estimadores utilizados en la simulación**

|  |  |
| --- | --- |
| **Discretas** | **Continuas** |
| * Media
* Varianza de máxima verosimilitud
* Varianza muestral
* Primer estadístico de orden
* Último estadístico de orden
 | * Media
* Mediana
* Varianza de máxima verosimilitud
* Varianza muestral
* Primer estadístico de orden
* Último estadístico de orden
 |

**Elaboración**: *R. Plúa*

En las distribuciones discretas con parámetros poblacionales estipulados por el investigador; por no ser factible la obtención de la mediana poblacional mediante una fórmula explícita que se pueda implementar en el programa, el estimador de la mediana no es trabajado. Para el caso de la distribución continua Normal Bivariada el estimador que se trabaja en la simulación es el del coeficiente de correlación. La función Variables utiliza a la vez a la subfunción Resultados.

 **3.3.2 Subfunción Resultados**

Estando creadas las variables de los parámetros de entrada para la población y el estimador seleccionado por el investigador, éstas son utilizadas para obtener los parámetros poblacionales, por ejemplo si el investigador selecciona la distribución exponencial y el estimador de la varianza muestral, con el parámetro poblacional ingresado podemos obtener el parámetro de la varianza poblacional. Está función a su vez define e inicializa las variables que se utilizarán para la presentación de los resultados, tales como las listas en las cuales se encuentran los estimadores, sesgo e intervalos de confianza obtenidos por el método Jacknife, los objetos “figura” que contienen los histogramas de las distribuciones para los estimadores simulados y las variables que contienen la varianza, sesgo de simetría, promedio, mínimo y máximo de los estimadores anteriormente mencionados. Esta función utiliza la subfunción Procedimiento\_Operaciones.

 **3.3.3 Subfunción Procedimiento\_Operaciones**

 Dependiendo del estimador seleccionado por el investigador esta función calcula las variables creadas en la función Resultados, es decir, genera 50 estimadores de los parámetros de la población estipulada, obtiene el sesgo de estimación y sus intervalos de confianza para cada uno de ellos. Una vez obtenido estos resultados calcula el sesgo de estimación promedio, la media de los estimadores, varianza, mínimo, máximo, sesgo de simetría y gráfica la distribución de frecuencia de los 50 estimadores, para cada método, Jacknife y Convencional. Aunque pareciera que está función realiza todo, no es así ya que a su vez está utiliza otras funciones como lo son Estimador\_Jacknife, Graficos y Generar\_Muestra.

 **3.3.4 Subfunción Generar\_Muestra**

Para generar los 50 estimadores en la función Procedimiento\_Operaciones, necesitamos generar una muestra aleatoria de tamaño n para a partir de esta calcular el estimador estipulado, por tanto en la función Generar\_Muestra se utilizan las variables que contienen los parámetros poblacionales ingresados por el investigador, para generar la muestra aleatoria de determinada población, en Matlab no es necesario realizar los algoritmos para la obtención de los números aleatorios de determinadas poblaciones, ya que trae incorporadas las funciones para la generación de números aleatorios en el módulo tools\statistical. Sin embargo para la generación de números aleatorios normales bivariados no existe una función que los proporcione por lo que se tuvo que realizar una función que los proporcione. A continuación presentamos los algoritmos para la generación de números aleatorios de distribuciones en las que no se puede obtener una fórmula explícita para la generación de los mismos.

**Binomial**

1. Se genera un número Ri.
2. Di Ri < pi entonces x=x+1 incrementamos i en1 y mientras i ≤ n retornamos al paso1.
3. X es el número de una población binomial.

**Binomial negativa con parámetros r y p.**

1. Se genera un número Ri.
2. Di Ri<pi entonces éxito = éxito + 1
3. Incrementamos i en1 y x=x+1, mientras éxito ≤ r retornamos al paso1.
4. X es el número de una población binomial.

**Hipergeométrica con parámetros N, k y n**

1. Si (k ≤ N) y (n ≤ N) entonces x=0 y r=0 ir al paso 2 caso contrario ir al paso 5.
2. r=r+1, se genera un número aleatorio u.
3. Si u <  entonces x=x+1
4. Mientras (r < n) and (x < N) ir al paso 2.
5. X es el número de una población Hipergeométrica.

**Poisson con parámetro λ**

1. Generamos un número aleatorio perteneciente a una distribución uniforme R.
2. Inicializamos: i=0, po=e-λ y F0=p0
3. Si R < Fi entonces x=i caso contrario incrementamos ,  y regresar al paso 3.

**Normal Multivariada con parámetros  y  covm (matriz positiva definida)**

* + - 1. Obtener la longitud del vector ****
			2. Obtener la factorización Cholesky de la matriz de covarianzas R=Cholesky(****)
			3. Generar una matriz Z nxd de números aleatorios normales estándar y un vector vnx1 de unos.
			4. X=Z (R) + v (****)’

 **3.3.5 Subfunción Estimador\_Jacknife**

 Esta función es llamada por la función resultados para la obtención del estimador Jacknife a partir de: un arreglo que contiene la muestra aleatoria que ha sido generada, y la variable donde se encuentra el estimador que el investigador seleccionó. Con estas variables puede aplicar el algoritmo del método Jacknife realizado en la sección anterior.

 **3.3.6 Subfunción Gráficos**

Esta función asigna al objeto “figura” los histogramas calculados a partir de los arreglos que contienen los 50 estimadores calculados por los métodos Jacknife y Convencional. Cabe recalcar que existen variables que se utilizan no en una sola función sino en muchas funciones del programa por tanto están son declaradas como globales como es el caso de aquellas que contienen los nombres del estimador y población seleccionados por el investigador, así como los tamaños muestrales con los cuales se generan las muestras aleatorias, entre otras.