**INFERENCIA ESTADÍSTICA**

**Control n°2**

4-abril-2019

**Profesor**: Francisco Javier Leiva

**Ayudante**: Sofía Muñoz

1. **Preguntas Abiertas (10pts. c/u)**
2. Explique que es el muestreo por estratos

El muestreo Estratificado pretende considerar distintas categorías que posean gran homogeneidad a su interior. Los estratos pueden ser según sexo, edad, regio, estado civil, etc. Lo importante es que todos los estratos de interés estarán bien representados en la muestra.

La distribución de la muestra se puede hacer tomando una muestra aleatoria simple de cada estrato. A la vez existe varias formas de combinar los resultados de la muestra, tales como las siguientes:

-Afijación simple

-Afijación proporcional

-Afijación optima.

1. Un muestreo aleatorio simple siempre será preferible como forma de realizar una muestra de una población en cualquier escenario.

Falso, a pesar de ser un método fácil de usar no es útil cuando tenemos poblaciones grandes. Este muestreo básicamente se trata de elegir los elementos de la población de uno en uno, de forma que cada individuo tenga la misma posibilidad de ser elegido.

Se puede realizar de distintas maneras:

* Sacar números de una bolsa
* Utilizar una tabla de números aleatorios
* Recurrir a números generados por algún programa

Dentro del muestreo aleatorio simple probabilistico tenemos, por ejemplo, muestreo estratificado, por conglomerado, muestreo sistemático, entre otros. En los no probabilistico están los muestreos por cuotas, bola de nieve, discrecional o subjetivo.

1. El número de observaciones que debe tener una muestra siempre será mayor mientras más holgura le queramos dar al estimador con respecto al parámetro.

Considerando: , mientras más holgura estemos dispuesto a aceptar (e) el tamaño de muestra necesitado es menor, por lo que sería falso, sin embargo, solo considerando "ceteris paribus" pues el objetivo es estimar bien, por lo que típicamente uno quiere menos diferencia entre el parámetro y el estimador.

Además, la muestra no necesariamente tiene que ser más grande, sino que debe ser representativa de la población. No se busca dar holgura, se busca estimar de forma eficiente e insesgada.

1. **Ejercicios (20pts.)**
2. (20pts.) Encuentre el tamaño muestral a utilizar si desea que el parámetro a estimar no sea diferente en más de 0,5 unidades al verdadero valor con un 90% de confianza. Considere que conoce el verdadero valor de la desviación estándar, y ésta es igual a *σ* = 3. Explique.

Sea e=0,5, y

El tamaño muestral es 98

1. **STATA (10pts.)**
2. Imagine que tiene ya “cargada” una base de datos en STATA, la que contiene una variable llamada <var1> y <var2>. Escriba la sintaxis para crear una variable la cual sea el doble de <var1> siempre que la <var2> sea igual a 15. Explique.

gen var = 2\*var1 if var2==15

Para crear una variable se usa el comando gen, en este caso multiplicamos var1 por 2 para que sea el doble y utilizamos if para denotar que de se consideren solo los valores igual a 15 con el signo igual (==)

1. Explique y escriba la sintaxis para poder generar un muestreo aleatorio simple en una base de datos ya “cargada” en STATA.

En una base de datos ya cargada, generamos una variable que se distribuya uniforme entre 0 y 1:

Generate var= uniform()

Luego ordenamos de menor a mayor la variable

Sort var

Luego se genera una variable que asigna valor 1 a la primera mitad de la base y el valor 2 a la segunda.

generate group = ceil(2 \* \_n/\_N)





