

Capítulo 1

Concepto de población, marco y muestra.

Muestreo probabilístico. Distribución de un estimador en el muestreo. Error cuadrático medio y sus componentes. Intervalos de confianza: Estimadores insesgados y sesgados. Métodos de selección. Probabilidad de la unidad de pertenecer a la muestra y propiedades. Comparación con el muestreo no probabilístico: Muestreo por cuotas.

1.1. Concepto de población, marco y muestra.

El muestreo consiste en seleccionar un subconjunto de una población al objeto de observarlo y como resultado inferir características de la población completa.

Cuando hablamos de muestreo, se suele presentar una **población** consistente en un número finito y conocido de individuos o unidades. Cada unidad tiene asociada una variable de interés, que es de la que se quiere obtener información, con un valor que se desconoce, pero que se supone fijo para cada unidad de la población. Por tanto no se trata de una variable aleatoria. Las unidades dentro de la población están identificadas y generalmente se numeran de forma correlativa.

Para obtener información sobre la variable de interés se selecciona un subconjunto o **muestra** de la población y se observa el valor de la variable para cada uno de los miembros del mismo asociado a la unidad a la que pertenece. Estos valores se utilizarán para deducir algún valor acumulativo de la característica observada.

El proceso consistente en seleccionar la muestra de la población y utilizarla para inferir algún tipo de información acerca de la característica observada se denomina muestreo. El objetivo del mismo es seleccionar una muestra lo más significativa posible de la población, y a partir de ella obtener un estimador de nuestra característica objeto de estudio lo más preciso posible.

Así, el conjunto de individuos acerca de los cuales deseamos hacer alguna inferencia recibe el nombre de **población objetivo**. Dado que en algunas circunstancias puede ocurrir que no tengamos acceso al total de la población para hacer nuestras observaciones, aparece el concepto de **población investigada**, que es el subconjunto de la población que tenemos la posibilidad de observar. A cada miembro individual de la población se le conoce como **unidad de la población**.

Si al elegir la muestra elegimos unidades de la población, diremos que estamos escogiendo **unidades elementales de muestreo**. También podemos escoger la muestra entre grupos no solapados de unidades elementales, en cuyo caso hablaremos de **unidades de muestreo compuestas**. En cualquier caso, las **unidades de muestreo** deben ser subconjuntos de la población disjuntos entre sí y cuya unión equivalga a toda la población a investigar.

A la lista ordenada de todas las unidades de muestreo se le da el nombre de **marco**. Lo ideal es que el marco tal que las unidades elementales que contenga coincidan con las que componen la población objetivo, pero en la mayoría de los casos esto no es así debido a desactualizaciones, errores y omisiones. En cualquier caso, debe estar lo suficientemente próximo a la población objetivo como para poder hacer inferencias acerca de la misma basándonos en la información obtenida del marco.

Al subconjunto de la población, compuesto por unidades de muestreo elegidas del marco, que se va a observar se le llama **muestra**.

1.2. Muestreo probabilístico.

Como hemos visto, el objeto del muestreo es inferir información de una población que no se puede observar completa por motivos de coste, incomodidad, imposibilidad logística, etc. por tanto, nos interesará elegir la muestra de forma que sea lo más representativa posible de la población y nos dé la mayor cantidad posible de información acerca de la misma, y con la menor cantidad de incertidumbre. Una forma de alcanzar este objetivo es seleccionar la muestra mediante un procedimiento aleatorio y aplicar los resultados de la teoría de la probabilidad y la inferencia estadística para hacer deducciones sobre la población. Esto es lo que se conoce como **muestreo probabilístico**.

Para que un procedimiento de muestreo se considere probabilístico debe cumplir las siguientes condiciones:

- Podemos definir el conjunto de muestras distintas que son posibles si el procedimiento se aplica a nuestra población de estudio.
- Cada muestra posible tiene asignada una probabilidad de ser elegida, que se conoce como **probabilidad de selección**. Como consecuencia, cada unidad de muestreo tiene asignada una probabilidad de pertenecer a la muestra.
-