



Teoría de Muestreo

Definiciones, tipos, aplicaciones.

Introducción

Fundamentación

- La lectura crítica de estudios de investigación incluye algunos interrogantes referidos a la delimitación de la población, la determinación de la muestra, si el tamaño elegido de la muestra es el adecuado y si el error de muestreo es apropiado.
- El muestreo es un elemento clave en la metodología de la investigación ya que implica seleccionar a un grupo de elementos que se utilizarán para dirigir un estudio.
- Es importante diseñar un plan de muestreo que defina el proceso de selección del grupo de elementos seleccionados.
- Es fundamental expresar claramente en todo trabajo de investigación los siguientes aspectos del diseño metodológico:
 - la población y las características que deben poseer los elementos para formar parte del estudio
 - el número de elementos que conforman la población, si este número se conoce con certeza o se puede estimar
 - el tipo de muestreo y la técnica utilizada
 - el tamaño de la muestra y el error de muestreo establecido.

Fundamentación (cont)

- La teoría del muestreo permite determinar de manera efectiva la muestra que refleje con exactitud las características de la población sometida al estudio, ya que no siempre es posible tomar a todos los elementos que conforman la población.
- Cuando se refiere a población en diferentes contextos de la investigación se mencionan los términos población objetivo, población investigada, población diana, población blanco, población accesible o *target group*.
- Estos términos no significan lo mismo, aunque algunos se utilizan como sinónimos, son distintas maneras de definir a la población de acuerdo al ámbito en el que se quiere estudiar (en ciencias sociales, ciencias biomédicas, ciencias económicas, marketing).

Definiciones

Universo

- Serie real o hipotética de elementos que comparten características definidas relacionadas con el problema de la investigación.

Población

- Metodológicamente, en un trabajo de investigación, se denota a la población como un conjunto definido, limitado y accesible del universo que forma el referente para la elección de la muestra. Es el grupo al que se intenta generalizar los resultados del estudio. Comprende todos los elementos (personas, familias, grupos, objetos, organizaciones, etc.) que presentan características comunes que se definen a través de criterios establecidos para el estudio.

Marco de muestreo

- Es el conjunto de las unidades de análisis delimitadas por los parámetros muestrales, aquellos que permiten realizar una selección asegurando la probabilidad de que todas sean elegidas.
- Existen los marcos de listas, donde los elementos o unidades muestrales son enumerados uno a uno.
- Ejemplos: listado de poseedores de tarjetas de crédito, abonados de una guía telefónica o empleados de una organización.

Muestra

- Es una colección de individuos extraídos de la población a partir de algún procedimiento específico para su estudio o medición directa. Una muestra es una fracción o segmento de una totalidad que constituye la población. La muestra es en cierta manera una réplica en miniatura de la población. Se estudian las muestras para describir a las poblaciones, ya que el estudio de muestras es más sencillo que el de la población completa, por que implica menor costo y demanda menos de tiempo.

Muestra representativa

- Para que una muestra sea representativa, tiene que contener las características relevantes de la población en las mismas proporciones en que están incluidas en tal población. Para evaluar la representatividad de la muestra se compara la media muestral con la media poblacional, si este parámetro se desconoce se puede estimar tratando de encontrar las medias obtenidas en trabajos anteriores que han analizado las mismas variables. Siempre hay un grado de error en las estimaciones lo que constituye el error muestral, una de las maneras de minimizarlo es diseñando un plan de muestreo adecuado.

Unidad de muestreo

- Unidad de análisis de la población
- Ej. Persona, hospital, laboratorio.

Fracción de muestreo

- Proporción de la población que forma parte de la muestra

Población objetivo

- Es el conjunto de la población sobre la cual vamos a obtener información.

Población muestreada

- Es el subconjunto de la población que contiene a los elementos que pueden ser elegidos para la muestra.
- Definida en el marco de muestreo.

Sesgo de selección

- **Selección de los grupos de estudio.** Por la forma como se seleccionan los grupos (por diferentes criterios de selección entre grupos) puede que éstos no sean comparables con todas las variables pronósticas. Esta comparabilidad es más difícil de asegurar en los estudios observacionales.
- **Pérdidas de seguimiento.** Cuando las personas que se pierden durante el seguimiento difieren de las que no lo hacen por variables relacionadas con la exposición y la enfermedad. La introducción de este sesgo está relacionada con los motivos de las pérdidas, más que con su número, aunque un número excesivo de pérdidas puede comprometer la potencia del estudio (poder estadístico) para detectar la relación / asociación estudiada. El sesgo relacionado con la pérdida de individuos durante el seguimiento es característico de los estudios de carácter prospectivo.
- **Supervivencia selectiva.** Este sesgo es importante en los estudios de casos y controles en los que se incluyen casos prevalentes, ya que éstos pueden ser los sobrevivientes de casos incidentes y la supervivencia puede estar relacionada con la exposición. Además, los casos prevalentes cambian sus hábitos en relación con la exposición y en consecuencia también se sesga la estimación del efecto.
- **Muestra no representativa de la población de estudio.** Esta situación puede ocurrir frecuentemente en los estudios transversales, aunque también puede presentarse en los estudios analíticos. Se introduce en el proceso de selección de la muestra procedente de la población de estudio.
- **Sesgo de detección.** Este sesgo se produce cuando el procedimiento utilizado para identificar el estado de enfermedad varía o se modifica según la presencia o ausencia de exposición.
- **Participación de voluntarios o efecto de autoselección.** Cuando existe una 'selección' por múltiples factores, algunos de ellos no medidos, y que pueden sesgar la estimación del efecto

Parámetro y estadístico

- **Parámetro:**
 - Medida numérica calculada con todos los elementos que componen una **población**.
 - Constante.
- **Estadístico:**
 - Medida numérica calculada con todos los elementos que componen una **muestra**.
 - Varía dependiendo de la muestra.

Diseño de muestreo

- Es un documento, plan o protocolo en el que se detalla el procedimiento de muestreo para el estudio de una población.
- Debe de contener toda la información sobre la población y metodología de muestreo, para garantizar el correcto análisis de los resultados.

Censo

Definición

- Algunas veces resulta útil examinar a todos los elementos de la población, en este caso se denomina censo.
- Como la realización de un censo demanda tiempo y recursos humanos, técnicos y económicos, sólo pueden hacerlo las grandes organizaciones como gobiernos, universidades u otros organismos nacionales e internacionales.
- Ejemplos: Censo de población y vivienda de los habitantes de la República Oriental del Uruguay. Censo nacional agropecuario. Censo de alumnos y docentes de la Universidad de la República.

Casos de usos de censos

- Construcción de un marco de muestreo.
- Cuando por el tamaño de la población los costos y tiempos necesarios no difieren significativamente de realizar un muestreo.
- Cuando para asegurar una muestra representativa es necesario un tamaño de muestra (n) muy grande, cercano al de la población (N). Esto ocurre generalmente en poblaciones con mucha variabilidad.
- Cuando es necesario precisión total.
- Cuando ya tenemos toda la información recopilada (de otros estudios o de registros médicos por ejemplo)

Inconvenientes del censo

- Lentos
- Caros, en parte por la cantidad de elementos de una población y también por la cantidad de personal capacitado necesario para realizarlos en un tiempo útil.
- Es necesario que el personal cuente con la capacitación y los tiempos adecuados para garantizar la integridad de los datos.
- No se pueden modificar luego de que están iniciados.

Muestreo

Solucionando los problemas del censo

Pero introduciendo problemas nuevos

Introducción

- La teoría del muestreo permite determinar de manera efectiva la muestra que refleje con exactitud las características de la población sometida al estudio, ya que no siempre es posible tomar a todos los elementos que conforman la población.

Casos de uso

- Pruebas destructivas.
 - Algunos tipos de estudios pueden requerir que para la medición del atributo el elemento pierda su integridad física.
(Si quiero probar una bala es necesario detonarla)
- Imposibilidad de realizar un censo.
 - Esto puede ser porque sea físicamente imposible estudiar a toda la población o porque los costos para hacerlo lo hagan prohibitivo.
- Uniformidad.
 - Menos encuestadores pueden estar mas coordinados, o es imposible disponer de mas personal entrenados, o los equipos necesarios son pocos, grandes o costosos.
- Oportunidad de la información.

Inferencia

- El muestreo da información completa de la muestra, pero nada sobre el resto de la población.
- A los procesos que nos permiten concluir características de una población a partir de datos de una muestra los llamamos inferencias estadísticas.
- Tiene un nivel de error.
- Si no tenemos en cuenta factores que modifican el nivel de error podemos llegar a conclusiones erróneas de una población.
- No obstante siempre se asume en el diseño de los estudios el riesgo de error tratando de minimizarlo desde la estructura del estudio. Esto se hace calculando el tamaño de muestra adecuado, el tipo de muestra o el tipo de inferencia que realizamos.

Distribución muestral

- Es la distribución de probabilidad del estadístico de estudio.
- Permite calcular como un estadístico tiende a variar a causa del muestreo aleatorio.
- La distribución es generada por la repetición del muestreo un gran número de veces.
- Esto permite estimar cuan probable es que el estadístico obtenido se acerque a la población, u obtener a partir del estadístico un rango de valores donde es más probable que el parámetro se encuentre

Tipos de muestreos

Muestreo no probabilístico

- La elección de los elementos no depende de la probabilidad sino de las causas relacionadas con las características de la investigación o de quien establece la muestra.
- Aquí el procedimiento no es mecánico, ni con base en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de una persona o de un grupo de personas y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación.

Muestreo no probabilístico (cont)

- Desde el enfoque cuantitativo y para determinado diseño, la utilidad de una muestra no probabilística reside no tanto en una “representatividad” de elementos, sino en una cuidadosa y controlada elección de sujetos con ciertas características definidas previamente en el planteamiento del problema.
- Es un muestreo que no permite generalizar ni tener certeza que la muestra sea representativa, pues no todos los miembros de la población tienen probabilidad conocida de ser elegidos
- Tipos:
 - Muestreo discrecional o de Conveniencia (elegido por el investigador)
 - Muestreo casual o incidental (voluntarios)

Muestreo aleatorio o probabilístico

- Las unidades de análisis tienen que estar dispuestas de modo tal que el proceso de selección dé una equiprobabilidad de selección a todas y cada una de las unidades que figuran en la población. Cada uno de los elementos de la población tiene la misma probabilidad de integrar parte de la muestra.

M.A.S.

Muestreo Aleatorio Simple

- Se utiliza cuando se conocen todos los elementos que conforman la población.
- Cada elemento de la población tiene la misma posibilidad de ser elegido para formar parte de la muestra.
- Se elabora una lista enumerativa de los elementos. Desde aquí se obtiene una serie de elementos que constituyen la muestra.
- El proceso consiste en trabajar con una tabla de dígitos al azar o una serie de números aleatorios generados en una planilla de cálculo y seleccionar a los que coinciden con el número al azar y el orden de la lista, o utilizando un programa estadístico que seleccione una cantidad de elementos aleatoriamente de una lista general que es la población.
- No se puede utilizar cuando hay alguna subpoblación con una marcada diferencia en la característica de estudio

M.S.

Muestreo Sistemático

- Los elementos se seleccionan de la población en un intervalo uniforme que se mide respecto de tiempo, orden o espacio.
- Se emplea si existe una lista ordenada de los elementos de la población o cuando se sabe cuántos elementos componen esa población.
- La técnica consiste en tomar cada k elementos de una lista que contiene todos los elementos de una población, eligiéndose al azar el primer elemento de la muestra.
- Para determinar el valor k se realiza el cociente entre el tamaño de la población N y el tamaño de la muestra, por ejemplo si se tiene una población de 150 personas y se desea tomar una muestra de 30 entonces el valor del intervalo será de 5

M.A.E.

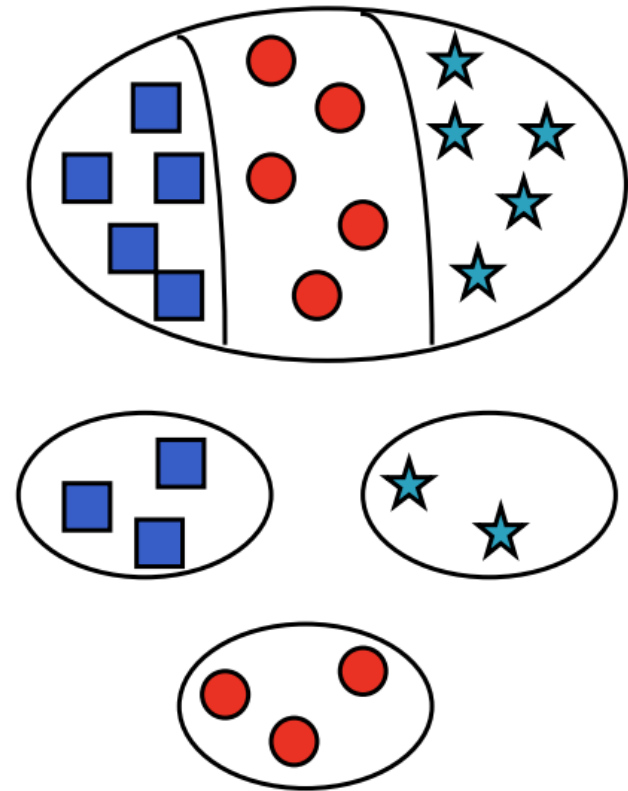
Muestreo Aleatorio Estratificado

- Es una variante del muestreo aleatorio simple. Se divide a la población en grupos homogéneos denominados estratos. Los elementos dentro de cada estrato se seleccionan aleatoriamente de acuerdo con una de las siguientes reglas:
 - Un número específico de elementos se extrae de cada estrato, y corresponde a la proporción de ese estrato en la población
 - Igual número de elementos se extraen de cada estrato, y los resultados son valorados de acuerdo con la porción del estrato de la población total.

M.A.E.

Muestreo Aleatorio Estratificado

- Esta técnica de muestreo resulta apropiada cuando la población ya está dividida en grupos, por que refleja de forma más precisa las características de la población y permite efectuar comparaciones entre los estratos conformados.
- Al extrapolar los resultados a la población hay que tener en cuenta el tamaño relativo del estrato con respecto al total de la población.
- Se busca que las unidades sean homogéneas dentro de los subgrupos y a su vez los estratos difieren entre sí de modo de aumentar la precisión de las estimaciones.
- Los estratos deben ser mutuamente excluyentes



M.C

Muestreo por Conglomerados

- Se utiliza cuando no es posible obtener una lista de todos los elementos de la población.
- Su empleo es adecuado si la población es muy grande y dispersa.
- Los conglomerados se caracterizan por ser homogéneos entre sí, pero internamente presentan un alto grado de heterogeneidad en sus componentes.
- La técnica consiste en lo siguiente: se divide a la población en grupos o racimos, luego se selecciona aleatoriamente algunos de esos grupos, por considerar que cada uno de ellos es representativo de la población y posteriormente se toma una muestra aleatoria de cada uno de los grupos que se han seleccionado.

M.C

Muestreo por Conglomerados

- Este procedimiento produce una muestra más precisa a un menor costo ya que se utiliza cuando hay variación considerable dentro de cada grupo, siendo los grupos similares entre sí.
- El conglomerado es común en los diseños polietápicos y en las muestras de zona geográfica.
- Cuando se muestrean conglomerados que contienen números de unidades desiguales, pueden utilizar el muestreo probabilístico proporcional al tamaño para que la probabilidad de selección del conglomerado sea igual a la proporción de unidades que contiene.



Muestreo Multietapa

- Muestreo en el que se procede por etapas: se obtiene una muestra de unidades primarias, más amplias que las siguientes; de cada unidad primaria se toman, para una submuestra, unidades secundarias, y así sucesivamente hasta llegar a las unidades últimas o más elementales.

Puede considerarse como una modificación del muestreo por conglomerados. Entonces, no forman parte de la muestra elementos o unidades de todos los conglomerados, sino que, una vez seleccionados los conglomerados aleatoriamente, se efectúan submuestras dentro de cada uno de ellos.

Muestreo con reemplazo

- Procedimiento de muestreo en el que los elementos se regresan a la población después de ser elegidos, de tal forma que algunos elementos de la población pueden aparecer en la muestra más de una vez.
- La probabilidad de selección es constante y si es una variable dicotómica sigue una distribución binomial.

Muestreo sin reemplazo

- En este caso cuando un elemento es seleccionado para la muestra no puede ser elegido nuevamente
- La probabilidad de selección se modifica con cada selección, y en una variable dicotómica sigue una distribución hipergeométrica.

Sesgo estadístico

- Error inherente a toda medición, medible, predecible y repetitivo.
- Se mide como la diferencia entre la esperanza de un estadístico y el parámetro que estima.
- También podemos ver la potencia de la estimación con la precisión de la predicción.

Teorema central del límite

- Es un teorema fundamental de probabilidad y estadística.
- Describe la distribución de la media de una muestra aleatoria proveniente de una población con varianza finita.
- Cuando el tamaño de la muestra es lo suficientemente grande, la distribución de las medias sigue aproximadamente una distribución normal.
- El teorema se aplica independientemente de la forma de la distribución de la población para muestras grandes ($n > 29$) y siempre para poblaciones con distribución normal.

Teorema central del límite

- La media de la distribución de las medias muestrales será igual a la media poblacional.
- La varianza de la distribución de las medias será igual a la varianza poblacional dividida por el tamaño de la muestra (n). Con los valores de desvío estándar se calcula dividiendo por la raíz cuadrada de n y se le llama error estándar.