

## Teoría Elemental del Muestreo



#### Muestreo Geológico

La muestra geológica es una unidad representativa de suelo, roca, mineral, fluido o gas que es seleccionada desde una larga masa o volumen para servir como un largo cuerpo o reflejar alguna característica específica o variación dentro de este. La simple racionalidad para Muestra (Barnes, 1981) es que, uno puede tomar un espécimen de casa, pero no de un afloramiento. Muestrear es tomar lo proveído de un espécimen para un propósito de clasificación y/o especial propósito (e.g., petrofabrica) análisis, ensayo, o testeo (incluyendo evaluación ingeniería geológica o geoquímico).

Porque la muestra, una vez extraída, toma el lugar del afloramiento en alguna forma, el prerrequisito básico de una muestra es que debe ser representativa de lo que sea que se muestree. La muestra representativa es dependiente, en gran parte, de la precisión geológica que precede al muestreo y el procedimiento usado para tomar la muestra.

"Lo malo del infinito es que es muy muy largo, especialmente la última parte."

Woody Allen (1935-)



## ¿Qué es el Muestreo?

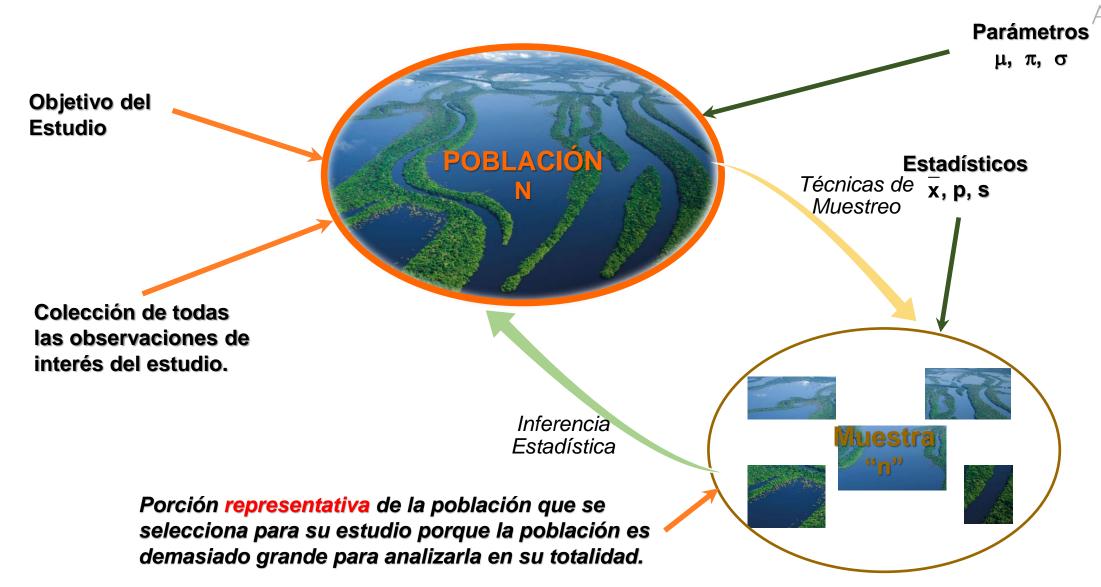


- El Muestreo es el método que se utiliza para diseñar y seleccionar una muestra.
- El muestreo es un proceso que consiste en extraer un conjunto de datos de un marco o de varios marcos de muestreo. El objetivo del muestreo es estimar los parámetros de la población ( $\mu$ ,  $\sigma$ ,  $\pi$ ) utilizando la información contenida en la muestra.
- La población puede ser finita, infinita numerable o infinita no numerable.
- El muestreo puede hacerse con o sin reemplazo. Si la población es finita pero muy grande, entonces es factible aplicar los conceptos de poblaciones infinitas.



## Población vs Muestra











- La imposibilidad física de revisar todos los integrantes de la población (casos).
- La naturaleza destructiva de algunas pruebas.
- Menor costo.
- Menor tiempo empleado y por tanto mayor rapidez en la recolección y análisis (para tomar decisiones oportunas).
- Datos disponibles limitados.
- Lo adecuado de los resultados de una muestra.





# Requisitos Básicos de una muestra

- La muestra debe ser representativa de la totalidad de la población en cuanto a características importantes para el problema en estudio. Se deben reflejar todos los elementos de la población en la muestra (*Muestra representativa*).
- El número de miembros seleccionados debe ser lo suficientemente para que las estimaciones tengan el grado de precisión adecuado (*Muestra adecuada*).
- El método de selección debe permitir el cálculo de una medida del error de muestreo.
- La muestra debe ser lo más eficiente posible en cuanto a tiempo y costo.
- El proceso de muestreo debe tener el mínimo de errores no muestrales.







- Establecimiento de Objetivos.
- Población Objetivo y Marco de Muestreo.
- Diseño de la muestra.
- Método de medición.
- Instrumento de medición.
- Capacitación de investigadores de campo.
- Prueba piloto.
- Organización del trabajo de campo.
- Análisis de los datos.





## **Población**



Conjunto completo de elementos, con alguna característica común, objeto de nuestro estudio.

## **Muestra**



Parte representativa de la población la cual nos ayudará a realizar inferencias acerca de esta, cumple con ciertas condiciones.





## Unidad de muestreo (grupo)

Ejemplo: Un testigo / una sección geológica / una quebrada





#### Unidad elemental

Ejemplo: Un fragmento rocoso / una caída de agua/ un mineral









- **Parámetro:** Es una medida que resume y describe información de una variable de la población. Describe el comportamiento de una característica de la población. Es una valor constante y para calcularlo se deben tener todos los datos de la población ( $\mu$ ,  $\pi$ ,  $\sigma$ ).
- Estadística o estadígrafo: Es una función de las n observaciones de una muestra aleatoria. Sirve de base los métodos estadísticos para hacer inferencias con relación a poblaciones y sus parámetros. La estadística se convierte en una variable porque su valor puede cambiar de una muestra a otra (ejemplo: la media muestral, la varianza muestral y la proporción muestral) y tienen distribución de probabilidad que se le llama distribución muestral.
- **Datos:** Son mediciones que son recolectadas de una población (o de una muestra).





- Variable: Característica de la población o muestra que se observa.
- Estadística Descriptiva: Proceso de recopilación, organización y presentación de un conjunto de datos.
- Estadística Inferencial: Utilización de un conjunto de datos muestrales para extraer inferencias o conclusiones sobre una población.
- Muestreo sin reemplazo: Cuando un elemento de la población es incluido en la muestra, ya no se puede seleccionar por segunda vez. Este tipo de muestreo se usa con mayor frecuencia.
- Muestreo con reemplazo: Cuando al seleccionar un elemento para la muestra, este se regresa a la población y por tanto se puede volver a seleccionar y, en consecuencia, puede aparecer más de una vez en la muestra.



## **El Marco Muestral**



Un marco muestral es una lista de las unidades a ser muestreadas (registros). Se espera que la diferencia entre el marco y la población sea pequeña.

## Ejemplos:

- Si tomamos un fragmento rocoso de 20cm^3 como unidad de muestreo de la Formación Vivian en la Cuenca Ucayali, todo el volumen de roca de la Formación Vivian en la Cuenca Ucayali es el marco muestral.
- El volumen de agua superficial y subterránea en la Subcuenca Camaná.

#### Definición de muestra:

Es una colección de unidades seleccionada de un marco o de varios marcos muestrales.

#### Observación:

La información de la muestra será tan buena como lo es la información contenida en el marco muestral.



## Error Muestral y Tamaño de la Muestra



Es el error natural que se origina por el hecho de trabajar con una muestra. Diferencia entre el estadístico de la muestra que sirve como estimación del parámetro desconocido de la población y el propio valor real del parámetro.

Por ejemplo (ejemplo imaginario), si se realizó una perforación para determinar el porcentaje de cobre en una Formación y los resultados mostraron que la ley era 61%. El análisis mostró que el error muestral era de más o menos 3%. Esto significa que el verdadero valor de la ley de cobre estaría entre 58% y 64%.

Las muestras más grandes proporcionan errores de muestreo más pequeños.

Un aumento en el tamaño de la muestra, será siempre más beneficioso en una muestra pequeña que en una grande, si se trata de reducir el error muestral.

El tamaño de la muestra se refiere siempre al número entero real de unidades extraídas de la población, y no a un porcentaje de la misma.

OJO: Mayor número de muestras es mejor siempre y cuando sean de calidad!!!



## Error no Muestral



Es aquel error ajeno al proceso de muestreo, difícil de acotar y motivado principalmente por una elevada tasa de no respuesta, errores de diseño de perforación, sesgos del geólogo recolector, etc.

Un control riguroso de todo el proceso de realización del estudio reduce este tipo de error que pueden restar gran parte de la validez de los resultados (principalmente debemos tener un flujo completo de trabajo desde la fase de precampo hasta la fase de entrega de resultados, en el muestreo considerar siempre el **QA/QC** para el estudio).

Normalmente, aumentan con el tamaño de la muestra. Recalcando, una muestra más grande no necesariamente significa que es mejor.

#### Ejemplos:

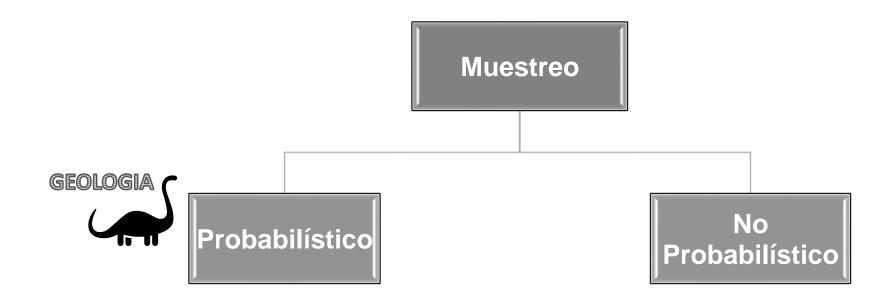
- Errores de los geólogos al tomar la muestra.
- Errores del precampo (malla de muestreo, planificación de muestreo).
- Errores en el instrumento (multiparámetro no calibrado, brújula no calibrada, entre otros).
- Errores en el trabajo de campo o ejecución.
- Errores en el marco muestral.



## ¿Cómo realizar el muestreo?



Existen diferentes modos de realizar el muestreo. Dependen de área de aplicación y de quien es el elemento o unidad de análisis. Por ejemplo, en estudios de mineralogía una porción de roca de 20 cm^3 es elemento donde el conjunto de éstos forman un cuerpo ígneo (la población). En estudios de geoquímica de aguas, una alícuota de 200ml de agua es el elemento y el conjunto de estas alícuotas forman el recurso hídrico en una cuenca (población). Existe múltiples aplicaciones dependiendo del campo en el que se realice el estudio aunque por lo general en la estadística tradicional se tienen los siguientes :







## MUESTREO PROBABILÍSTICO

El muestreo probabilístico se caracteriza porque cada elemento de la población tiene una probabilidad conocida y no nula de ser seleccionada.

La selección se hace por un proceso de aleatorio y el valor de la probabilidad se determina de acuerdo a la naturaleza del diseño muestral. Generalmente se considera la equiprobabilidad como base (todos los individuos tienen la misma probabilidad de ser elegidos para formar parte de una muestra).

Los diferentes muestreos probabilísticos son los siguientes:

- Muestreo Aleatorio Simple (MAS)
- Muestro Estratificado (ME)
- Muestreo Sistemático (MS)
- Muestreo por Conglomerados (MC)
- Muestreo Polietápico (MP)



# Muestreo Aleatorio Simple



Es el método más sencillo de selección de los elementos de una población y es el que más se utiliza.

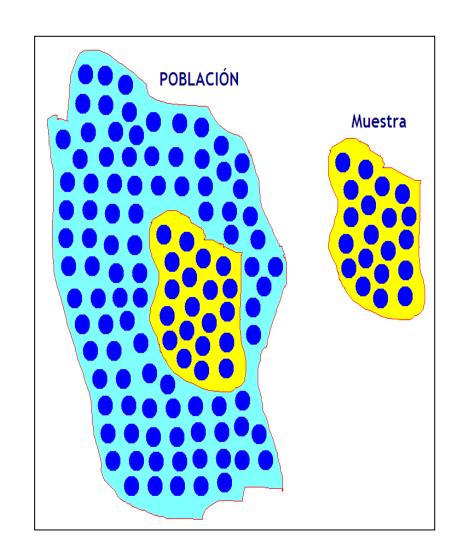
La selección aleatoria simple de una población se efectúa en forma tal que cada muestra posible de un tamaño dado tiene la misma probabilidad de selección. Es decir que todos los elementos que componen la población tienen igual probabilidad.

Este tipo de muestreo probabilístico, se basa fundamentalmente en que la población es homogénea. Y es allí donde radica su gran deficiencia, puesto muy pocas poblaciones son homogéneas.

Los investigadores a campo suelen seleccionar un valor al azar que distinga latitud y otro que distinga longitud para cada punto.

Se selecciona con algún medio mecánico (bolillas dentro de una bolsa, tabla de números aleatorios, números aleatorios generados en una computadora).

Finalmente debemos mencionar que tienen poca o nula utilidad práctica cuando la población que estamos manejando es muy grande o cuando es imposible enumerar todos sus elementos.







# Muestreo Aleatorio Simple

## **EJEMPLO**

Si se desea elegir una muestra de 50 rocas a partir de una población de 500 rocas de un museo geológico. La muestra se elige a partir de una lista de las rocas en los archivos del museo, a cada uno de los cuales se les asigna un número de identificación consecutivo.

Formas de elegir la muestra por el muestreo aleatorio simple:

• Poner en una urna 500 papeles con la numeración correlativa de 1 a 500 y seleccionar en forma aleatoria 50 papeles. Los números seleccionados corresponderán a los 50 rocas que conformarán la muestra (puesto que cada roca posee un número de identificación).



## **Muestreo Sistemático**

Cuando se tiene una población de N elementos y se necesita una muestra de tamaño "n"; en éste caso, lo que se hace es ordenarlos en función de una variable y después se elige en forma aleatoria un elemento entre los k = N/n primeros, y luego se elige de manera sistemática el que esté k lugares después del primer elemento, y así sucesivamente.

Esto significa que cada k elementos de la población se considera uno como miembro de la muestra. El número i de inicializador será un número al azar entre 1 y k.

El problema en este tipo de muestreo está en los casos que se dan periodicidades en la población ya que al elegir a los miembros de la muestra con una periodicidad constante (k) podemos introducir una homogeneidad que no se da en la población.



Selección	Orden	Código	Nota
	1	201229	06
	2	201244	06
1	3	201245	06
	4	201294	06
	5	201298	06
	6	201232	07
	7	201246	07
	8	201252	08
2	9	201258	08
	10	201267	08
	11	201270	08
	12	201271	08
	13	201271	08
	14	201223	09
3	15	201236	09
	16	201253	09
	17	201273	09
	18	201229	10
	19	201262	10
	20	201279	10
4	21	201241	11
	22	201217	12
	23	201237	12
	24	201285	13
	25	201257	14
	26	201260	14
5	27	201233	15
	28	201233	16
	29	201242	16
	30	201248	17

n = 5

k = 6

a = 3



# Muestreo Estratificado



Por lo general se emplea cuando la población es heterogénea pero contiene determinados grupos homogéneos. Un estrato se define como un conjunto de unidades homogéneas entre sí y son heterogéneas con las unidades de otros estratos. Primero se divide a los elementos de la población en grupos llamados estratos. Cada elemento de la población debe pertenecer a un sólo estrato. Después de formar los estratos se toma una muestra aleatoria de cada uno de ellos (afijación proporcional o no proporcional). La distribución de la muestra en función de los diferentes estratos se denomina *afijación*, y puede ser de diferentes tipos:

Afijación Simple: A cada estrato le corresponde igual número de elementos muestrales.

**Afijación Proporcional:** La distribución se hace de acuerdo con el peso (tamaño) de la población para cada estrato (Por ejemplo si un estrato posee el 20% de la población, contribuirá con el 20% de la muestra).

Afijación Óptima: Se tiene en cuenta la dispersión de los resultados, es decir, a mayor variabilidad en la población, se debe tomar mayor número de elementos de ella para que sea representativa la muestra.

Una afijación es un concepto teórico interesante, pero tiene poca aplicación ya que no suele conocerse la variabilidad en la población normalmente.

Si los elementos de cada estrato son homogéneos, el muestreo estratificado producirá, empleando un menor tamaño muestral, resultados tan precisos como el muestreo aleatorio simple.

Reduce el costo del muestreo al reducir los tamaños de muestra sin perder precisión.

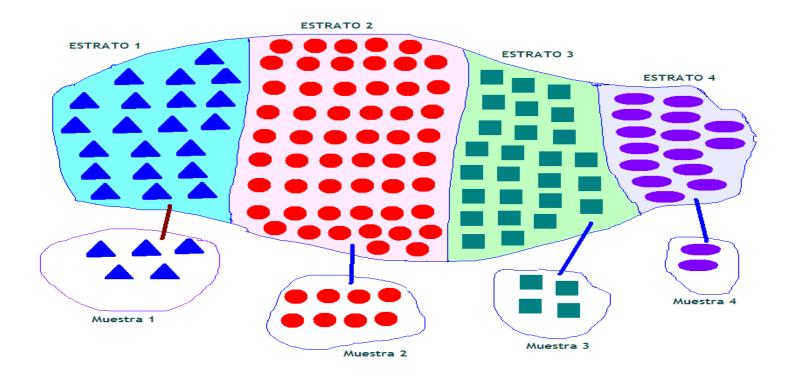
El criterio para la formación de los estratos queda a discreción del investigador (por ubicación, edad, tipo de análisis (agua, roca, suelo, petróleo, entre otros), tipo de Formación, escala de trabajo, magnitud del proyecto, etc.).

Se requiere información para estratificar la población.



# Muestreo Estratificado





#### Criterios en el Muestreo Estratificado

- \* Homogeneidad (dentro): Entre elementos de un mismo estrato.
- Heterogeneidad (entre): Entre estratos.
- Independencia: Entre estratos al seleccionar la muestra



# Muestreo Estratificado



#### EJEMPLO CASO AFIJACIÓN PROPORCIONAL:

Se desea estudiar la cantidad de Cobre en superficie de determinadas Formaciones de la Cuenca Marañon. Para ello de caracteriza a cada Formación (estratos) y se decide obtener una muestra proporcional en área que cubren con un muestreo estratificado; es decir se requiere que en la muestra seleccionada el número de elementos de cada estrato tenga la misma proporción que se encuentra en la población. Debemos previamente tener el "mapa geológico" a la escala que determinaremos trabajar. Se determino tomar 1000 muestras en toda la Cuenca (con un método de selección geológico con criterios estadísticos para el tamaño de muestra).

Si se emplea el muestreo aleatorio simple se corre el riesgo de no seleccionar ningún elemento de una Formación Condorsinga que cubra poca área Superficial.

Estrato	Formación	Área Total	%	Número de muestra por Formación
1	Vivian	150km^2	8.06	80.6
2	Aramachay	200km^2	10.75	107.5
3	Condorsinga	10 km^2	0.54	54
4	Cushabatay	180km^2	9.68	96.8
5	Cachiyacu	150km^2	8.06	80.6
6	Contaya	350km942	18.82	188.2
7	Cabanillas	500km^2	26.88	268.8
8	Chonta	320km^2	17.21	172.1

¿Cómo elijo el tamaño de muestra?

¿Dónde tomo la muestra?

¿Qué otros criterios son necesarios para que sea suficiente y representativa?



## Muestreo por Conglomerados



Se divide a los elementos de la población en conjuntos o grupos separados denominados conglomerados, unidades primarias o racimos.

Cada elemento de la población debe pertenecer a un sólo conglomerado. Luego se toma una muestra aleatoria simple de los conglomerados.

Conforman la muestra todos los elementos dentro de cada conglomerado muestreado. La medición se realiza a todos los elementos del conglomerado seleccionado.

Cada uno de los conglomerados puede tener diferente número de elementos y deben ser homogéneos pero los elementos dentro de los conglomerados deben ser heterogéneos. En el caso ideal, cada conglomerado es una versión representativa, en pequeña escala, de toda la población.

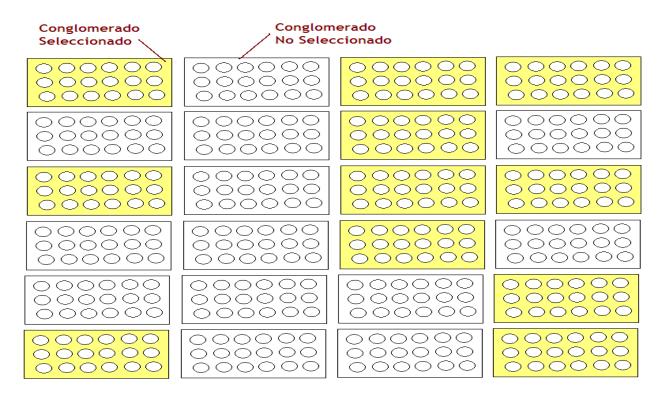
Una de las principales aplicaciones del muestreo por Conglomerados es el muestreo de áreas (cuando la población se encuentra dispersa en un área geográfica grande), en las que los conglomerados son áreas bien definidas.

Debemos seleccionar un número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y en investigar después todos los elementos pertenecientes a los conglomerados elegidos.



# Muestreo por conglomerados





## Criterios en el Muestreo por Conglomerados

- Homogeneidad (dentro): Entre elementos de diferentes conglomerados.
- Heterogeneidad (entre): Entre elementos del mismo conglomerado



# Muestreo por Conglomerados



## **EJEMPLO:**

Se desea obtener información sobre el contenido en roca aflorante de Cobre en la Formación Cazadores de la Cuenca Marañón. Con el propósito de subdividir la cuenca en áreas menores, se divide dicha en subcuencas; es decir se forman 20 conglomerados (subcuencas) y se eligen 10 cuencas a fin de tomar datos a toda la Formación en estos conglomerados. La muestra total estaría dada por los afloramientos rocosos de la Formación Cazadores en las subcuencas.

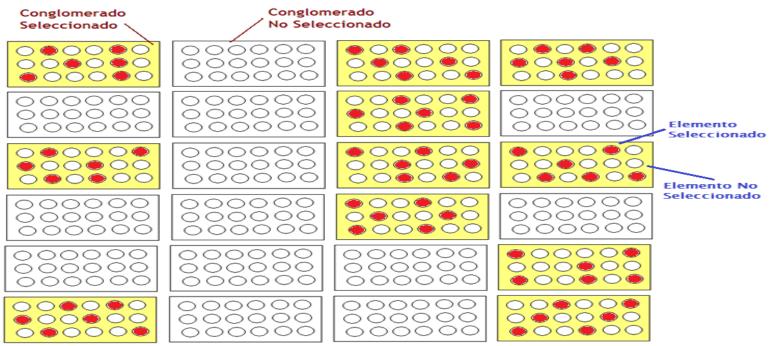
Por lo general se requiere un tamaño muestral total mayor que el muestreo aleatorio o el muestreo estratificado. Sin embargo, se origina ahorro porque cuando se envía al geólogo a tomar muestra a todos los elementos del conglomerado (subcuencas en el ejemplo) se podrán obtener muchas observaciones muestrales en un tiempo relativamente corto y con mayor facilidad.



# Muestreo polietápico



Se combina el muestreo aleatorio simple o sistemático con el muestreo por conglomerados: Primero se realiza un muestreo por conglomerados; segundo, se elige una muestra aleatoria. Dicha muestra puede ser obtenida por muestreo aleatorio simple o estratificado o sistemático. En ese sentido, se puede decir que se ha utilizado dos 2 etapas de muestreo. Y claro está, es posible tener más de 2 etapas.





# **MUESTREO** no PROBABILÍSTICO



Se llama muestreo no probabilístico a cualquier método de muestreo en que la oportunidad de escoger un elemento determinado de la población es desconocida, independientemente de que existan o no otras especificaciones.

Generalmente para estudios exploratorios, el muestro probabilístico resulta excesivamente costos y se acude a métodos no probabilísticos, aún siendo consciente que **no se puede hacer generalizaciones** (inferencia estadística), pues no se tiene la certeza que la muestra extraída sea representativa, ya que no todos los elementos de la población tienen la misma probabilidad de ser elegidos. En general se seleccionan a la muestra siguiendo determinados criterios procurando que la muestra sea representativa, los resultados se obtenidos son solo descripciones de esa muestra.





## Los diferentes muestreos no probabilísticos son los siguientes:

- Muestreo Casual o Fortuito.
- Muestreo de Selección Experta o Muestreo a Juicio.
- Muestreo por Cuotas.
- Muestreo Intencional.
- Muestreo de Bola de Nieve.
- Muestreo de Poblaciones Móviles.



# MUESTREO no PROBABILÍSTICO



- Muestreo Casual o Fortuito los elementos de estas muestras son elegidos simplemente porque son accesibles, pueden ser sujetos voluntarios o porque son personas que saben expresarse, etc. Al preguntar a los comuneros en zonas mineras de donde extraen el mineral.
- Muestreo de Selección Experta o Muestreo a Juicio es utilizado para seleccionar especímenes, unidades, o porciones, representativas o típicas.
- Muestreo por Cuotas es una forma de muestreo no aleatorio ampliamente utilizado en encuestas de opinión, de mercado y otras semejantes. Se asienta generalmente sobre la base de un buen conocimiento de los estratos de la población o de los individuos más "representativos" o "adecuados" para los fines de investigación. Los encuestadores reciben la instrucción de obtener cuotas especificas a partir de las cuales se construye una muestra relativamente proporcional a la población, con respecto a unas cuantas variables demográficas. Dentro de las cuotas, se supone que los encuestadores obtienen individuos representativos. La naturaleza de los controles y de las instrucciones dependen del juicio experto de la persona que practica este tipo de muestreo.



# MUESTREO no PROBABILÍSTICO



- Muestreo Intencional se caracteriza por un esfuerzo deliberado de obtener muestras "representativas" mediante la inclusión en la muestra de grupos supuestamente típicos. Es muy frecuente su utilización en sondeos en donde ha existido una tendencia generalizada anterior.
- Muestreo Bola de Nieve consiste en seleccionar una muestra inicial o básica de individuos y establecer en cada entrevista qué nuevas personas de la población en estudio han de entrevistarse, para así integrar la muestra completa. En un sentido muy amplio, la primera muestra puede seleccionarse en forma intencional o estar constituida por voluntarios.
- Muestreo de Poblaciones Móviles en este caso se depende de los métodos de captura-marca-recaptura. La población total se estima de la proporción de los individuos en la recaptura (insectos, peces, venados), que han sido capturados previamente y marcados. Hay modelos teóricos ingeniosos que se utilizan para expresar en forma explícita las suposiciones del método.

# TODO LO REVISADO ES UNA INTRODUCCIÓN BÁSICA A LO QUE RESPECTA AL MUESTREO PARA INFORMACIÓN MÁS EXACTA REVISAR LAS SIGUIENTES FUENTES:

#### **BIBLIOGRAFÍA:**

#### **Medio Ambiente y Geotecnia:**

https://www.epa.gov/sites/production/files/2015-06/documents/g5s-final.pdf



https://www.researchgate.net/publication/284308665 Probabilistic and Statistical Methods in Engineering Geology I Problem Statement and Introduction to Solution

#### Estadística Intuitiva (Sociales):

https://www.springer.com/gp/book/9783319560700#otherversion=9783319560724

#### **BIBLIOGRAFÍA:**



#### Estadística en Geología (curso general):

http://maps.unomaha.edu/Maher/GEOL2300/week1/week1.html

#### Water Resource in Geology:

https://www.usgs.gov/mission-areas/water-resources/science/national-water-quality-project-sampling-methods?qt-science\_center\_objects=0#qt-science\_center\_objects
https://www.usgs.gov/mission-areas/water-resources/science/national-field-manual-collection-water-quality-data-nfm?qt-science\_center\_objects=0#qt-science\_center\_objects
https://www.usgs.gov/special-topic/water-science-school/science/water-quality-sampling-techniques?qt-science\_center\_objects=0#qt-science\_center\_objects

#### Mining Ore in Geology:

https://clu-in.org/conf/tio/r10hardrock3\_030513/Lecture-3-Kathy-Smith\_ExpandedVersion.pdf

https://core.ac.uk/download/pdf/145046601.pdf

https://www.mdpi.com/2075-163X/8/6/232/pdf

https://dggs.alaska.gov/webpubs/usbm/b/text/b356.pdf

https://library.seg.org/doi/10.1190/segam2017-17666908.1

https://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0370-44672014000300005



"LO QUE ESCUCHO LO OLVIDO. LO QUE VEO LO RECUERDO. PERO LO QUE HAGO, LO ENTIENDO." AUTOR: ANÓNIMO

**VAMOS A RESOLVER UN EJERCICIO** 



## **BIBLIOGRAFÍA ESTADÍSTICA APLICADA A GEOLOGÍA EXTRA:**

Analytics AoZ

Practical Methods for Data Analysis (US EPA QA/G-9, 2000)

Helsel, D. R., & Hirsch, R. M. (2002). *Statistical methods in water resources* (Vol. 323). Reston, VA: US Geological Survey.

Salvador Figu eras, M y Gargallo, P. (2003): "Análisis Exploratorio de Datos, 5campus.com, Estadística <a href="http://www.5campus.com/leccion/aed">http://www.5campus.com/leccion/aed</a>>

Ramalle-Gómara, E., & De Llano, J. A. (2003). Utilización de métodos robustos en la estadística inferencial. *Atención Primaria*, 32(3), 177-182.

Verzani, J. (2005). Using R for introductory statistics. CRC press.

Cohen, Y., & Cohen, J. Y. (2008). Statistics and Data with R: An applied approach through examples. John Wiley & Sons.

Arnaldo Mangeaud (2014). Estadística aplicada a las Ciencias Geológicas. Universidad nacional de Córdova.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, *Statistical methods in water resources: U.S. Geological Survey Techniques and Methods*, book 4, chapter A3, 458 p.





D. Mikhnev and E. V. Tregubova .Estimation of Erros of Sampling, Sample Preparation, and Sample Analysis in the Monitoring Scheme of Hydrochemistry Shop of the Achinsk Alumina Refinery. Russian Journal of Non-Ferrous Metals, 2007, Vol. 48, No. 1, pp. 6–9.

DENNIS R. HELSEL and TIMOTHY A. COHN .Estimation of Descriptive Statistics for Multiply Censored Water Quality Data. Water Resources Division, U.S. Geological Survey, Reston, Virginia. 1998