

UNIDAD 2: MÉTODOS DE RECOLECCIÓN DE DATOS

Estadística
Lic. en Nutrición
UCEL

CENSO

Una Licenciada en Nutrición quiere conocer cuál es la prevalencia de sobrepeso en la población adulta (mayores de 18 años).

¿Cómo podemos conocer este valor?

Una opción es realizar un censo.

Un **censo** es una muestra formada por todos los elementos que componen la población.

CENSO

¿Qué problemas se te ocurren si emprendés un censo de los adultos de Rosario?

- Costo
- Tiempo
- Imposible si hay que destruir una unidad para estudiar una característica de la misma

En argentina, el INDEC (Instituto Nacional de Estadísticas y Censos) realiza un censo de población aproximadamente cada 10 años. También se realizan otros censos con orientaciones específicas (por ejemplo, el censo económico).

ALGUNAS DEFINICIONES

Población: conjunto total de objetos o individuos que presentan características comunes observables, definidas en un cierto tiempo lugar. Su tamaño es indicado con la letra **N**.

Unidad: objeto o persona individual, se la denomina sujeto si la población consiste en personas.

Muestra: parte de la población que es usada para lograr información. Su tamaño es indicado con la letra **n**.

ALGUNAS DEFINICIONES

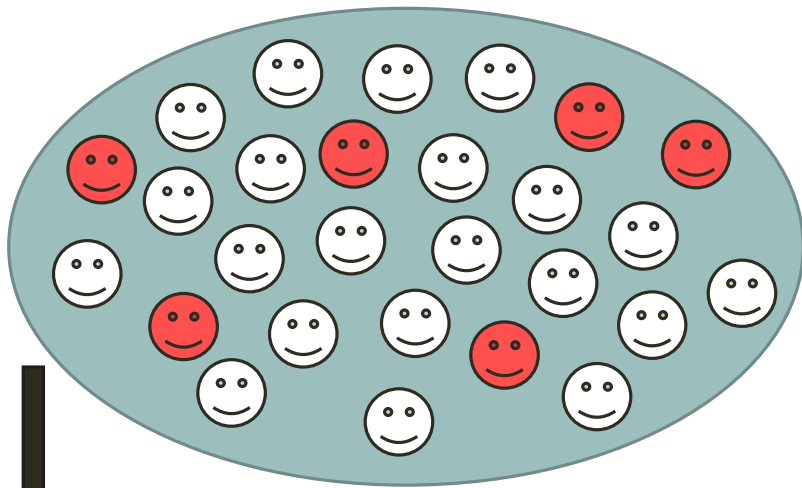
Variable: característica de interés a ser medida en cada unidad de la muestra.

Parámetro: medida resumen que se calcula a partir de todas las unidades de la población. Ejemplos: p , μ . Es un valor fijo.

Estadística: medida resumen que se calcula usando todas las unidades de la muestra de la población. Es variable, ya que toma valores distintos en diferentes muestras.

ALGUNAS DEFINICIONES

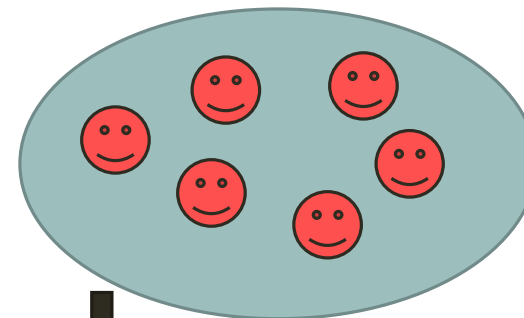
POBLACIÓN



PARÁMETROS

μ, p, σ

MUESTRA

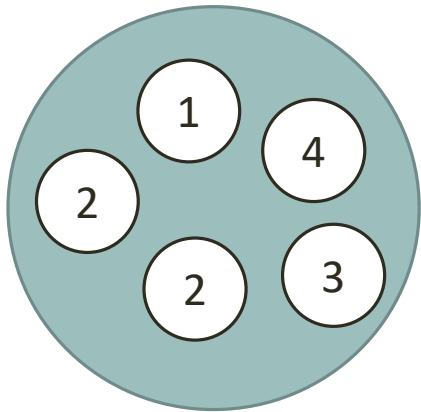


ESTADÍSTICAS

\bar{x}, \hat{p}, s

PARÁMETRO VS ESTADÍSTICA

Supongamos una población de 5 personas adultas, para la cual se quiere estudiar la cantidad de frutas que comen por día.



El parámetro de interés es:

- μ : el número promedio de frutas consumidas por día en la población

$$\mu = \frac{1 + 2 + 2 + 3 + 4}{5} = \frac{12}{5} = 2,4$$

PARÁMETRO VS ESTADÍSTICA

Sacamos una muestra de tamaño $n = 2$ y obtenemos las personas que comen 1 y 2 frutas. La estadística es:

- \bar{x} : el número promedio de frutas consumidas por día en la muestra

$$\bar{x} = \frac{1 + 2}{2} = \frac{3}{2} = 1,5$$

Ahora sacamos una muestra de tamaño $n = 2$ y obtenemos las personas que comen 2 y 4 frutas. La estadística es:

$$\bar{x} = \frac{2 + 4}{2} = \frac{6}{2} = 3$$

Vemos que el valor de la estadística no necesariamente coincide con el valor del parámetro, y que varía según la muestra seleccionada.

MANOS A LA OBRA N° 1



SOBREPESO EN NIÑOS DE MENOS DE 5 AÑOS

El 10% de los niños menores de 5 años en Argentina tienen sobrepeso. En una muestra aleatoria de 400 niños de menos de 5 años de esa población, se encontró que el 12% presentan sobrepeso.

- 1) ¿Cuál es la población?
- 2) ¿Cuál es la muestra?
- 3) ¿Cuál es la variable en estudio?
- 4) Identifique y defina el parámetro. ¿Cuál es su valor?
- 5) Identifique y defina la estadística. ¿Cuál es su valor?

MANOS A LA OBRA N° 1



¿PARÁMETRO O ESTADÍSTICA?

Un extracto de “Health Update: Aerobics hieren los oídos” indica:

Médicos del hospital Henry Ford en Detroit, estudiaron 125 clases de aerobics en 5 clubes de salud, y encontraron que los niveles de sonido de la música en el 60% de las clases excedía el límite de seguridad.

Los resultados establecidos en este párrafo están basados en u estudio de 125 clases de aerobics en 5 clubes de salud (no todas las clases de aerobics en todos los clubes de salud).

Este 60% es una _____ y el tamaño de la muestra es $n =$ _____.

La población de interés es _____
y el parámetro es _____.

¿TODAS LAS MUESTRAS SON BUENAS?

La Licenciada en Nutrición que quiere conocer cuál es la prevalencia de sobrepeso en la población adulta (mayores de 18 años) decide tomar como muestra a los pacientes que atiende en su consultorio durante el mes de marzo de 2020.

¿Es esta una “buena” muestra? ¿Qué problemas te parece que puede tener?

CUIDADO! No cualquier parte de la población representa una “buena” muestra.

SESGO

Un método de muestreo es **sesgado** si produce resultados que difieren sistemáticamente de los verdaderos de la población.

Un programa de TV llevó a cabo la siguiente encuesta de opinión:

“¿Cree que la obesidad es un problema serio? ¡Permítanos conocer su respuesta! ¡Esta noche!

- *Si contesta SÍ llame al 1-900-446-6444*
- *Si contesta NO llame al 1-900-446-6445*

La tarifa es 50 centavos el primer minuto.”

¿Quiénes tendrán interés en responder? ¿Hay alguna traba que puede impedir la participación de ciertos individuos?

¿Considera los resultados de esta encuesta de opiniones fidedignos? Explique.

SESGO

- **Sesgo de selección:** es la tendencia sistemática sobre el procedimiento de muestreo para excluir o incluir cierto tipo de unidades.
- **Sesgo de no respuesta:** es la distorsión que se logra cuando un gran número de unidades seleccionadas para la muestra no responden o se niegan a responder y esas no respuestas tienen una tendencia a ser distintas de los que responden.
- **Sesgo de respuesta:** es la distorsión que se logra por la forma de pregunta o el comportamiento de entrevistador puede afectar la respuesta.

MUESTRAS CONVENIENTES Y VOLUNTARIAS

Una **muestra conveniente** es una muestra que de aquellas unidades que se consiguen fácilmente.

Muestras voluntarias o convencionales son generalmente sesgadas.

MANOS A LA OBRA N° 2



TAMAÑO DE LA FAMILIA (Para resolver 2.3)

Se llevó a cabo un estudio para estimar el tamaño promedio de las casas de familia (hogares) en USA. Un total de 1000 personas fueron aleatoriamente seleccionadas e interrogadas respecto del número de personas que viven en sus casas. El promedio de estas 1000 respuestas fue 4,6.

- 1) ¿Cuál es la población de interés?
- 2) ¿Cuál es la variable en estudio?
- 3) ¿Cuál es el parámetro en estudio?

MANOS A LA OBRA N° 2



PREGUNTA SENSIBLE

Un departamento educativo realizó un relevamiento deseando conocer el consumo de bebidas alcohólicas entre sus estudiantes. Una de las preguntas realizadas fue : “¿Cuántos vasos de bebidas alcohólicas consume por semana?”. A esta pregunta respondió aproximadamente el 30% de los alumnos encuestados.

- 1) ¿Qué tipo de sesgo puede ocurrir en este estudio?
- 2) ¿Cómo piensa que será el consumo semanal promedio de bebidas alcohólicas de todos los alumnos en comparación con el consumo promedio de los alumnos que respondieron la pregunta en cuestión?

MUESTREO PROBABILÍSTICO

Para evitar los problemas ilustrados anteriormente se puede usar algún método de muestreo probabilístico.

Con estos métodos podemos cuantificar la precisión de usar una estadística muestral para estimar el parámetro poblacional correspondiente.

Un método de muestreo que asigna a cada unidad de la población una chance no negativa de ser seleccionada se denomina **método de muestreo probabilístico**.

MUESTREO PROBABILÍSTICO

Estudiaremos:

- Muestreo simple al azar
- Muestreo estratificado
- Muestreo sistemático
- Muestreo por conglomerados.

MUESTREO SIMPLE AL AZAR

Una **muestra simple al azar** de tamaño n es una muestra de n unidades seleccionadas de alguna manera tal que **cada muestra posible de tamaño n tenga la misma chance de ser seleccionada** que cualquier otra muestra de tamaño n .

Muestras de distinto tamaño tienen distintas chances de ser seleccionadas.

El método es imparcial: toda unidad tiene la misma chance de participar en la muestra.

MUESTREO SIMPLE AL AZAR

Para poder extraer una muestra utilizando el método de Muestreo Simple al Azar es necesario contar con un listado de todos los individuos que conforman la población de interés.

- **Marco de población:** lista de todos los elementos o individuos de los cuales se extraerá la muestra.

Esta lista sirve como **población objetivo**, y generalmente es distinta de la población real.

MANOS A LA OBRA N° 3



SITUACIÓN NUTRICIONAL. Aprendemos como usar una Tabla de Números Aleatorios

Un grupo de investigadores realiza un estudio sobre el estado nutricional de una cohorte de ingresantes a la carrera de Lic. En Nutrición. Dicha cohorte está formada por $N=80$ alumnos. Debido a que el cuestionario que se plantea realizar es muy extenso, los investigadores decidieron seleccionar 10 ingresantes al azar y entrevistarlos.

- 1) Utilice la fila 10 de la tabla de números aleatorios para seleccionar una muestra simple al azar de tamaño $n=5$ de esta población.
- 2) Uno de los objetivos del estudio es determinar la proporción de alumnos celíacos en la cohorte. En la muestra seleccionada se encontró que uno de los alumnos era celíaco. Estime el parámetro de interés.

MANOS A LA OBRA N° 3



PARA PENSAR DE LA PÁGINA 95

¿Cuándo resulta fácil seleccionar una muestra simple al azar? ¿Siempre es posible seleccionar este tipo de muestras?

¿Cuándo resulta difícil seleccionar una m.s.a.? ¿Por qué?

DESVENTAJAS DEL MSA

Un Licenciado en Nutrición quiere saber cuál es el consumo promedio de azúcares de sus pacientes. El 20% de sus pacientes son diabéticos, por lo que es sabido que deberían evitar o al menos restringir lo más posible el consumo de azúcar.

Extrae una muestra aleatoria simple de 30 pacientes. Ninguno de los pacientes en la muestra es diabético.

¿Qué problema te parece que puede tener la muestra seleccionada? ¿El Licenciado estimará correctamente el consumo promedio de azúcar de todos sus pacientes?

Hay situaciones donde el MSA puede resultar poco eficiente o inadecuado.

MUESTREO ESTRATIFICADO

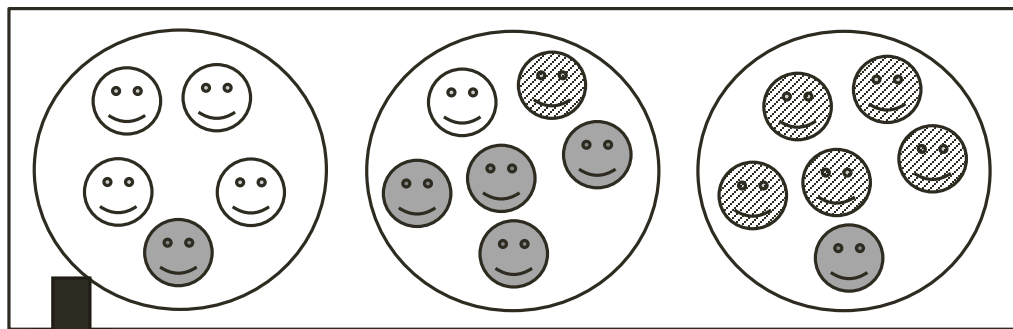
Una **muestra aleatoria estratificada** se obtiene dividiendo o estratificando la población en **subgrupos** o **estratos mutuamente excluyentes** y sacando una **MSA** de unidades **de cada estrato**.

- **Subgrupos mutuamente excluyentes:** cada unidad de la población pertenecerá a un solo estrato.

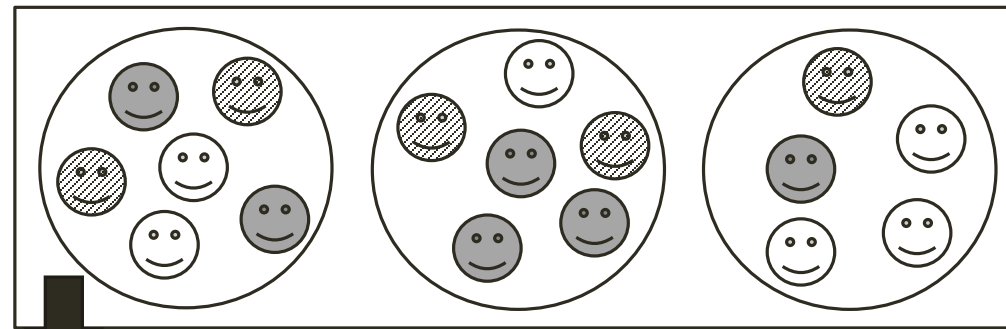
La idea detrás de la división en estratos es que las unidades que compartan estrato sean muy similares entre sí, mientras que las unidades de diferentes estratos difieran lo más posible. Es decir, **que la variabilidad dentro del estrato sea pequeña, mientras que la variabilidad entre estratos sea grande.**

MUESTREO ESTRATIFICADO

Este método es más efectivo cuando las unidades dentro del estrato individual son muy semejantes con respecto a la característica que se mide y los estratos son muy diferentes entre sí.



Es conveniente usar muestreo estratificado

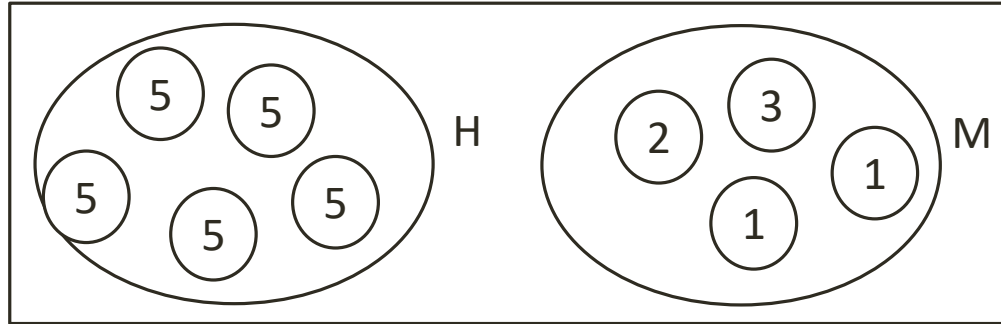


No es conveniente usar muestreo estratificado

En esos casos, proporciona estimaciones más precisas que las que se obtendrían con el MSA. Es **insesgado** cuando se usan estimadores **ponderados**.

MUESTREO ESTRATIFICADO

Supongamos que estamos estudiando la cantidad de porciones de carne que consumen las personas en una población que tiene 2 estratos, hombres y mujeres:



Si se quiere extraer una muestra aleatoria de tamaño 4, ¿cuántas unidades conviene seleccionar de cada estrato?

MANOS A LA OBRA N° 4

LICENCIAS PARA COMEDORES ESCOLARES



Una ciudad tiene otorgadas 40 licencias a comedores escolares. Un inspector de esta ciudad debe verificar que dichos comedores cumplan con las ordenanzas municipales de sanidad, y para ello inspeccionará 12 comedores por mes. Estratificará los comedores según la cantidad de alumnos a cargo que tengan las escuelas:

- Un comedor que alimenta a menos de 50 niños es pequeño. El estrato I está compuesto por 10 comedores escolares.
- Un comedor que alimenta a 50 o más niños es grande. El estrato II está compuesto por 30 comedores escolares.

Si no hay información previa sobre la variabilidad de los estratos puede muestrearse en proporción al tamaño del estrato y el tamaño de la población.

Seleccione las unidades que conforman la muestra. Para el estrato I use la fila 20 y la columna 11. Para el estrato II use la fila 30 y la columna 6.

MANOS A LA OBRA N° 4



PARA PENSAR DE LA PÁGINA 100

- 1) ¿Cuándo se toma un tamaño de muestra más grande en un estrato que en otro?
- 2) Cuando forma los estratos, ¿cómo debe ser la variabilidad dentro de cada estrato comparada con la variabilidad entre los estratos?
- 3) Una muestra estratificada, ¿es una muestra simple al azar? Explique

MUESTREO SISTEMÁTICO

Una Lic. En Nutrición que trabaja en un sanatorio decide investigar cuál es la proporción de los pacientes internados requieren una interconsulta con el sector de Nutrición. Para estimar este valor, decide estudiar las historias clínicas de los pacientes internados en el último año, y registrar cuáles de ellas sugieren que el paciente se hubiera beneficiado de una interconsulta. La lista de los paciente internados durante el último año está ordenada por fecha y hora de ingreso.

¿Le parece práctico generar una M.S.A.?

En muchos casos realizar un MSA o estratificado no es posible o resulta muy poco práctico o costoso. En algunos casos se puede aprovechar el ordenamiento presente en el listado de individuos para elegir la muestra.

MUESTREO SISTEMÁTICO

Para una **muestra sistemática** 1 en K ordenamos las unidades poblacionales de alguna manera y las dividimos en grupos de K unidades. Seleccionamos aleatoriamente 1 de las K primeras unidades de dicha lista ordenada, la unidad “k”. Esta unidad seleccionada es la primera a ser incluida en la muestra. Continuamos seleccionando cada k-ésima unidad de los grupos de K elementos del listado hasta completar el tamaño de la muestra deseado.

MUESTREO SISTEMÁTICO

Supongamos una población de 19 unidades, rotuladas de la A a la S, y queremos tomar una muestra sistemática de 1 en 4.

A	B	C	D		E	F	G	H		I	J	K	L		M	N	O	P		Q	R	S
1	2	3	4		5	6	7	8		9	10	11	12		13	14	15	16		17	18	19

En este caso, $K=4$. Veamos que el tamaño muestral varía según el valor de k que sea elegido aleatoriamente. El valor k es un número entre 1 y 4.

Supongamos que $k=3$, entonces la primera unidad seleccionada en la muestra es C, la segunda es G (posición $k+K$), la tercera es K ($k+K+K$), y así sucesivamente.

MANOS A LA OBRA N° 5



PARA PENSAR DE LA PÁGINA 105

Consideremos nuevamente las 19 unidades que componen la población, rotuladas de la A a la S. Se plantea tomar una muestra sistemática de 1 en 4.

- 1) Genere la muestra utilizando la fila 2 de la tabla de números aleatorios.
- 2) ¿Cuál es el tamaño de la muestra anterior?
- 3) Genere la muestra utilizando la fila 4 de la tabla de números aleatorios.
- 4) ¿Cuál es el tamaño de la muestra anterior?
- 5) El tamaño de la muestra, ¿es fijo? Explique.

MUESTREO DE CONGLOMERADOS

En muchos casos realizar un muestreo estratificado no es posible o resulta muy poco práctico o costoso.

Algunas veces, las poblaciones están organizadas en grupos de unidades o sujetos: familias, escuelas, hospitales. Otras veces pueden definirse grupos de manera artificial por su cercanía (geográfica o de otro tipo).

En esta situaciones, resulta más económico y practicable elegir varios grupos y estudiar a todas sus unidades, que elegir individuos de cualquier grupo (usando cualquiera de los tipos de muestreo estudiados).

MUESTREO DE CONGLOMERADOS

En el **muestreo de conglomerados** las unidades poblacionales se agrupan en conglomerados o clusters. Uno o más conglomerados (clusters) son seleccionados aleatoriamente. Si un cluster es seleccionado, todas las unidades serán incluidas en la muestra.

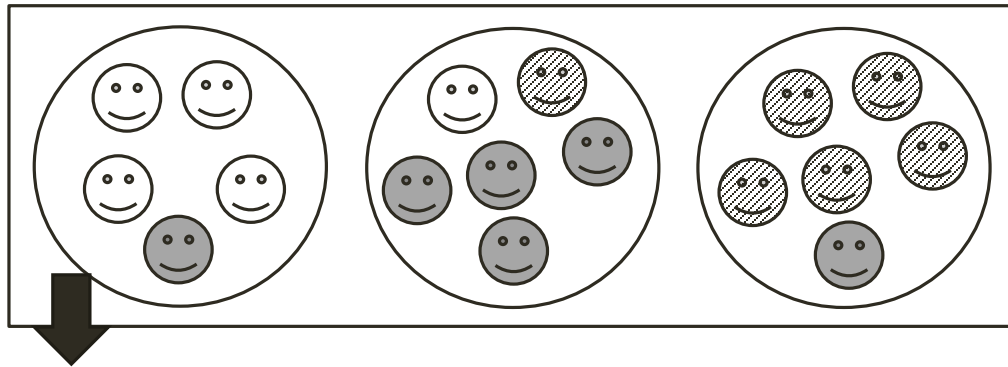
Algunos conglomerados están naturalmente definidos, como las familias, los hospitales, las escuelas, etc.

Como desventaja, a veces los clusters que están definidos naturalmente no describen adecuadamente la variabilidad presente en los estratos de la población.

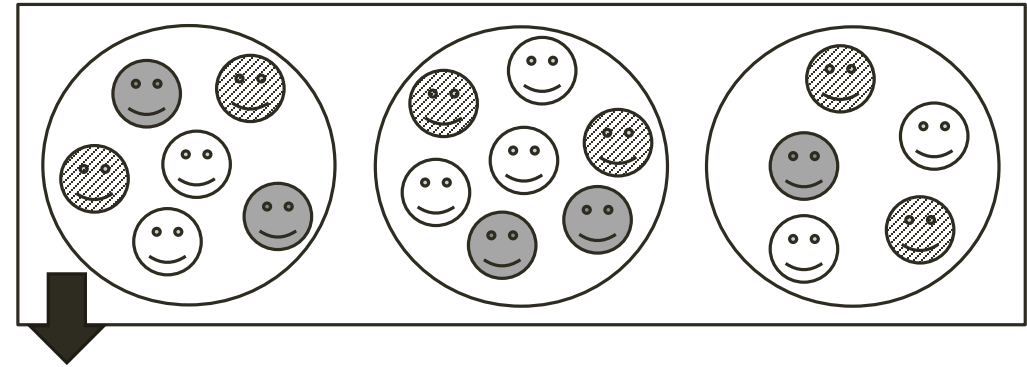
MUESTREO DE CONGLOMERADOS

En el muestreo estratificado los estratos están formados de tal manera que las unidades dentro del estrato son homogéneas.

En el muestreo de conglomerados, los conglomerados están formados como una “mini población”. Las unidades dentro de un conglomerado son heterogéneas.



Es conveniente usar muestreo estratificado



Es conveniente usar muestreo de conglomerados

MANOS A LA OBRA N° 6



PARA PENSAR DE LA PÁGINA 111

- 1) ¿Una muestra por conglomerados es una muestra estratificada? Justificar.
- 2) Cuando formamos los conglomerados, ¿cómo debería ser la variabilidad de las unidades dentro de cada conglomerado comparada con la variabilidad entre conglomerados?
- 3) ¿Este criterio es el mismo que para el muestreo estratificado?

MANOS A LA OBRA N° 7



PARA RESOLVER 2.12 DE LA PÁGINA 111. Mencione el método de muestreo

- 1) Un cargamento de 1000 bolsas de 2kg de papa ha llegado a un comedor escolar. Este cargamento fue embarcado en 50 cajas con 20 bolsas cada una. De las 50 cajas, 5 fueron seleccionadas aleatoriamente y se obtuvo el peso promedio de las 100 bolsas de papas que contenían.
- 2) Un miembro de una facultad desea tomar una muestra de 1600 alumnos de una carrera. Cada alumno tiene un código o un número identificador. Se dispone del listado de todos los números identificatorios. El profesor selecciona un número aleatorio entre los 16 primeros de dicho listado y después elige los otros integrantes de la muestra a intervalos de 16 números identificatorios de la lista.
- 3) El decano desea tomar una muestra de los 1600 alumnos de una carrera. Decide entrevistar a los 100 primeros alumnos que ingresen a un mismo curso el próximo lunes a la mañana.

MANOS A LA OBRA N° 7



PARA RESOLVER 2.13 DE LA PÁGINA 112. ¿Qué método de muestreo podría haber usado?

Una población consta de 12 personas. Estas personas están listadas según la tabla:

Población	Estrato I	Estrato II
Conglomerado I	Nora(1) Gabriel(2)	Marta(3) Alberto(4)
Conglomerado II	Augusto(5) Eugenio(6)	Gastón(7) Dante(8)
Conglomerado III	Lucía(9) Pablito(10)	Germán(11) Memi(12)

Para cada muestra de tamaño 4, indique cuál método la generó. Ayuda: es posible que una misma muestra pueda ser resultado de varios métodos.

Muestra	Método de muestreo		
	Simple al azar	Estratificado	Conglomerado
a) Augusto, Eugenio, Gastón, Dante			
b) Nora, Pablito, Augusto, Gabriel			
c) Augusto, Eugenio, Marta, Memi			