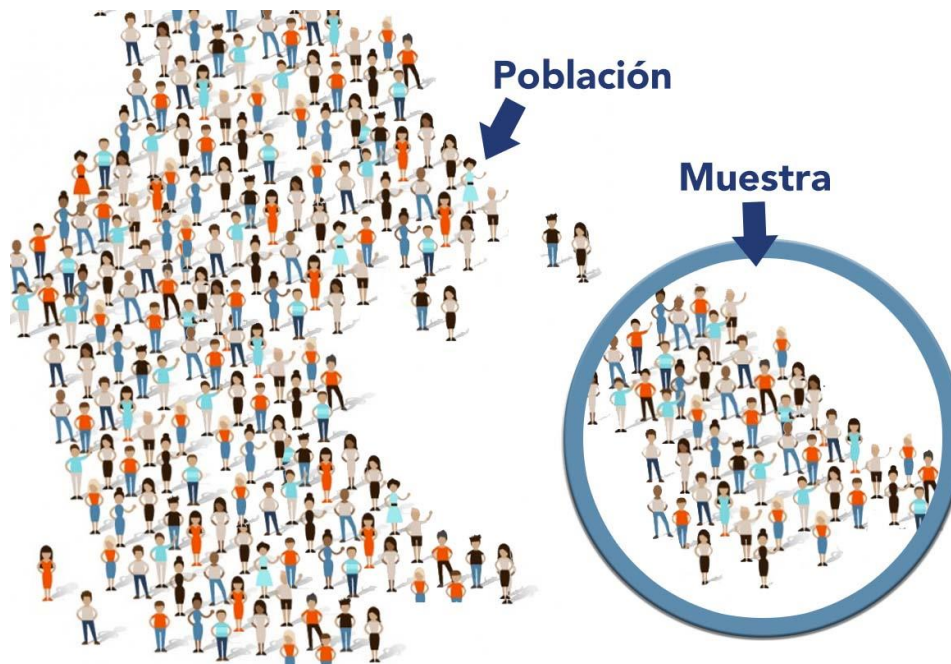


Muestreo

Técnica estadística que presenta métodos para seleccionar de manera **óptima y representativa** un subconjunto de elementos de la población para estudiarlos.



¿Que condiciones debe satisfacer una buena muestra?



1. Poseer un tamaño considerable.
2. Representar todas las características de variabilidad presentes en la población.

¿Por qué muestrear la población?

A menudo no es posible estudiar la población completa, algunas de las principales razones por la que es necesario muestrear son:

1. La naturaleza destructiva de ciertas pruebas.
2. La imposibilidad física de revisar todos los integrantes de la población.
3. El costo de estudiar todos los integrantes de una población.
4. Tiempo disponible para encuestar a toda la población.
5. Los resultados de la muestra son adecuados.

¿Por qué muestrear la población?



Cuchara para catar la sopa
a diario



Olla para la sopa a diario



Cuchara para catar la sopa
con invitados



Olla para la sopa con invitados

“El tamaño de muestra no es proporcional al tamaño de la población”

¿Por qué muestrear la población?



Cuchara para catar la sopa
con invitados



Cuchara para catar la sopa
a diario



Olla para la sopa a diario



Olla para la sopa con invitados

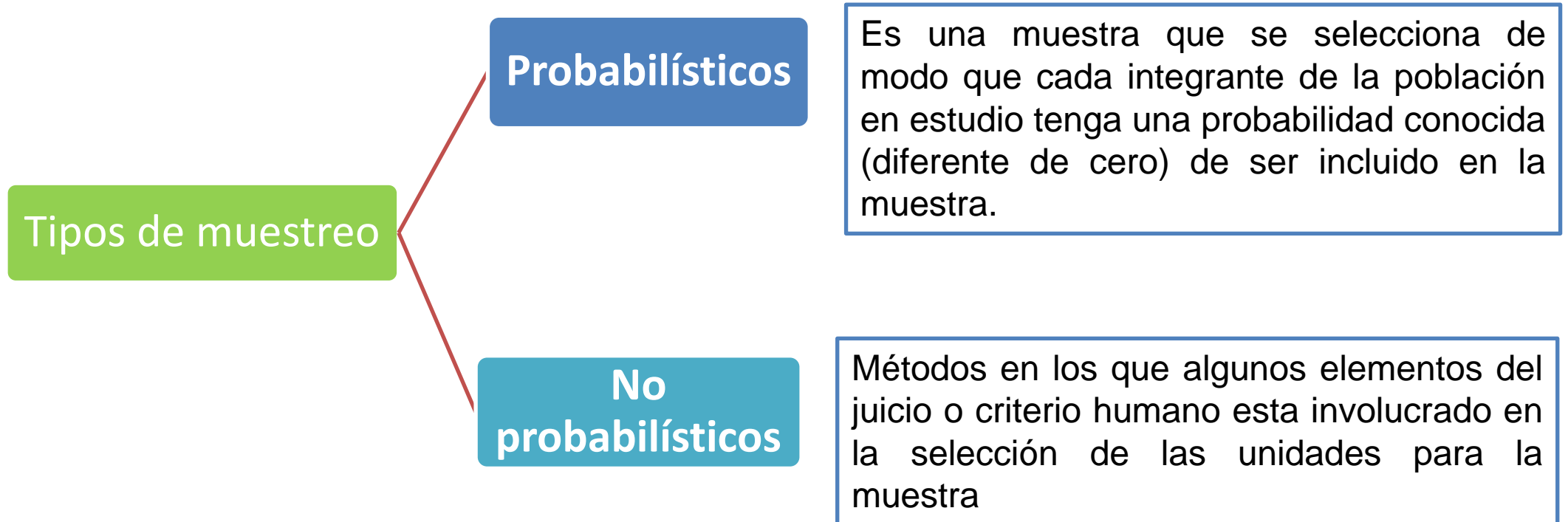
Doctor, ¿para saber mi grupo sanguíneo me va usted a sacar una muestra con el 10% de mi sangre?



¿Para saber en promedio cuantas patas tienen los cangrejos del Pacífico, necesito capturar el 10%?

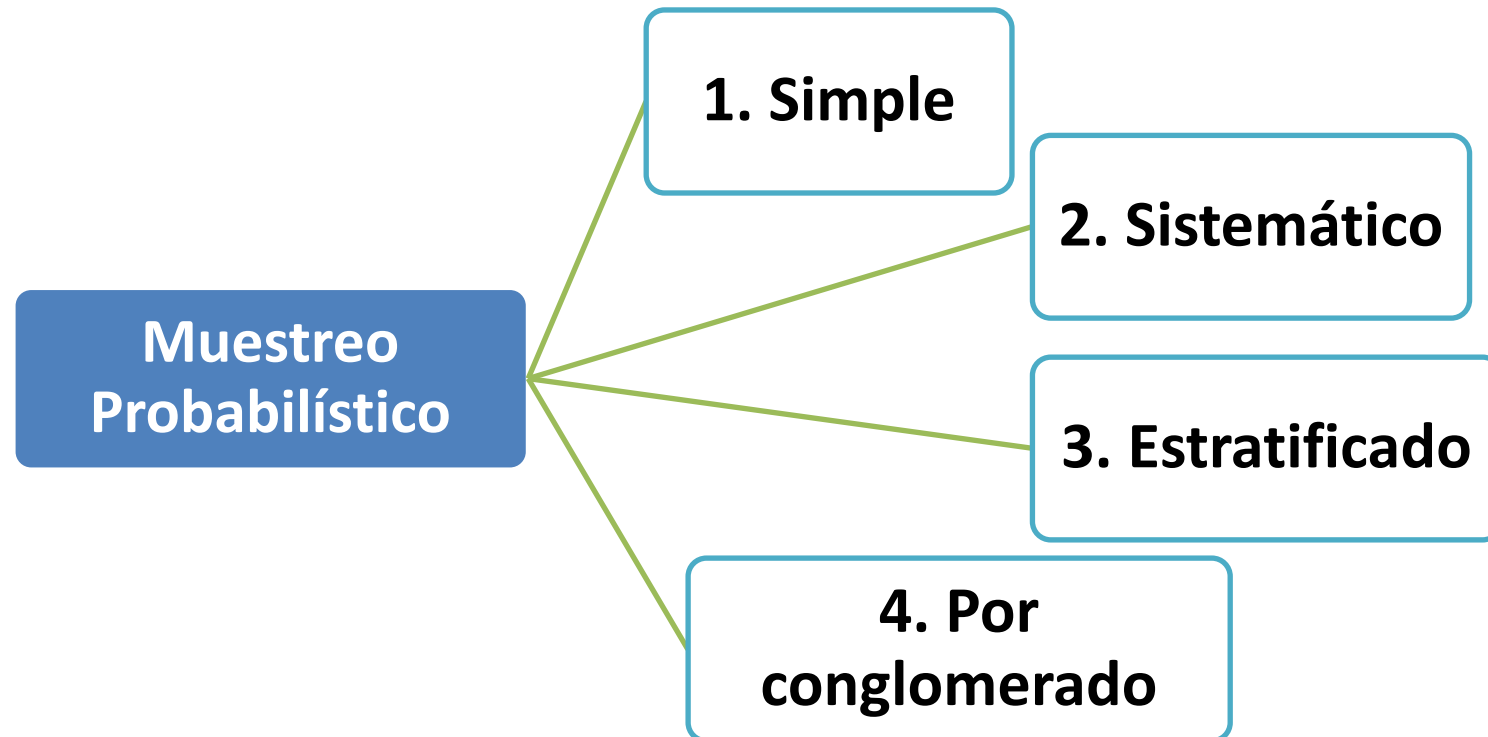


Tipos de Muestreo



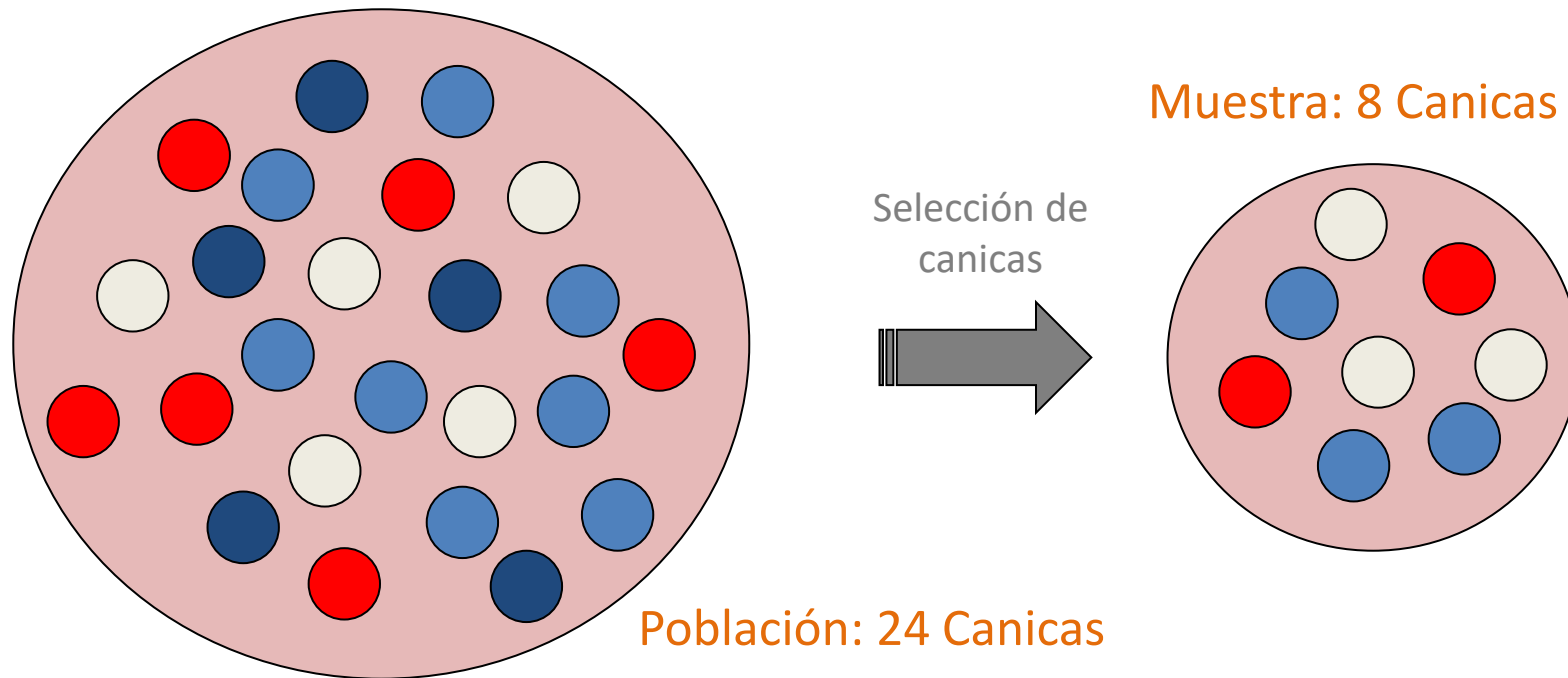
Muestreo Probabilístico

Los métodos probabilísticos de muestreo están soportados por las estructuras del azar donde se elimina o mitiga la subjetividad que podría influir en la elección de las unidades que integrarán la muestra y se podrá medir el grado de precisión de las estimaciones.



Muestreo Aleatorio Simple

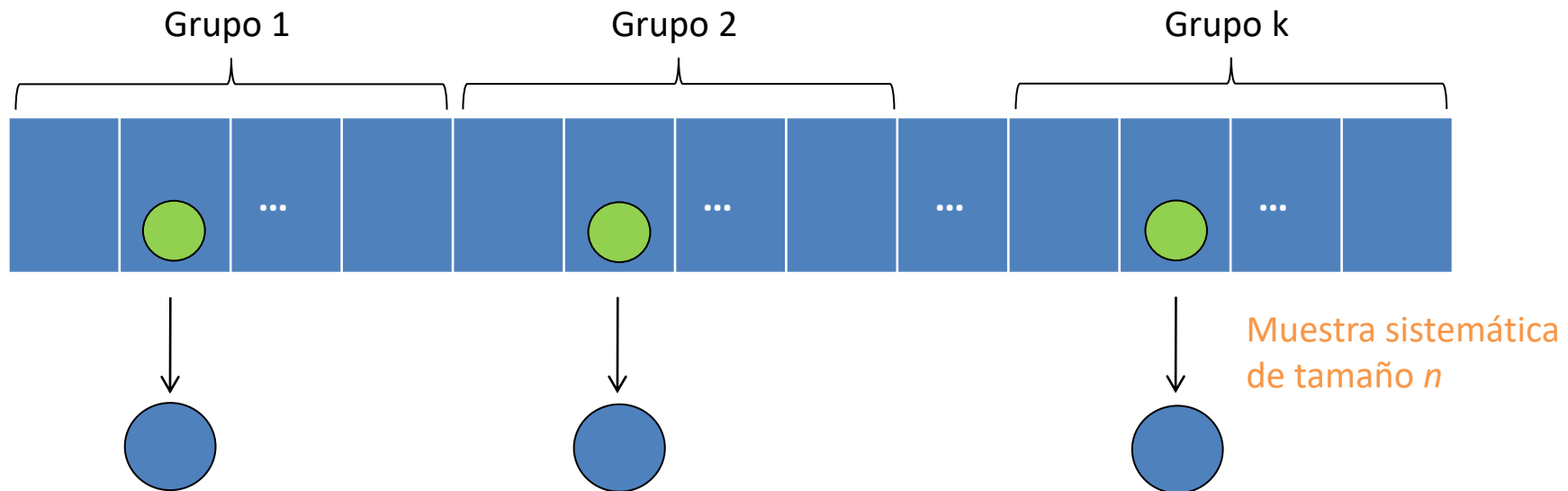
Las unidades son escogidas individual y directamente por un proceso aleatorio en el que cada unidad tiene la misma probabilidad de ser seleccionada como cualquier otra unidad.



Este tipo de muestreo asume que la población es relativamente homogénea respecto a la característica a estudiar.

Muestreo Sistemático

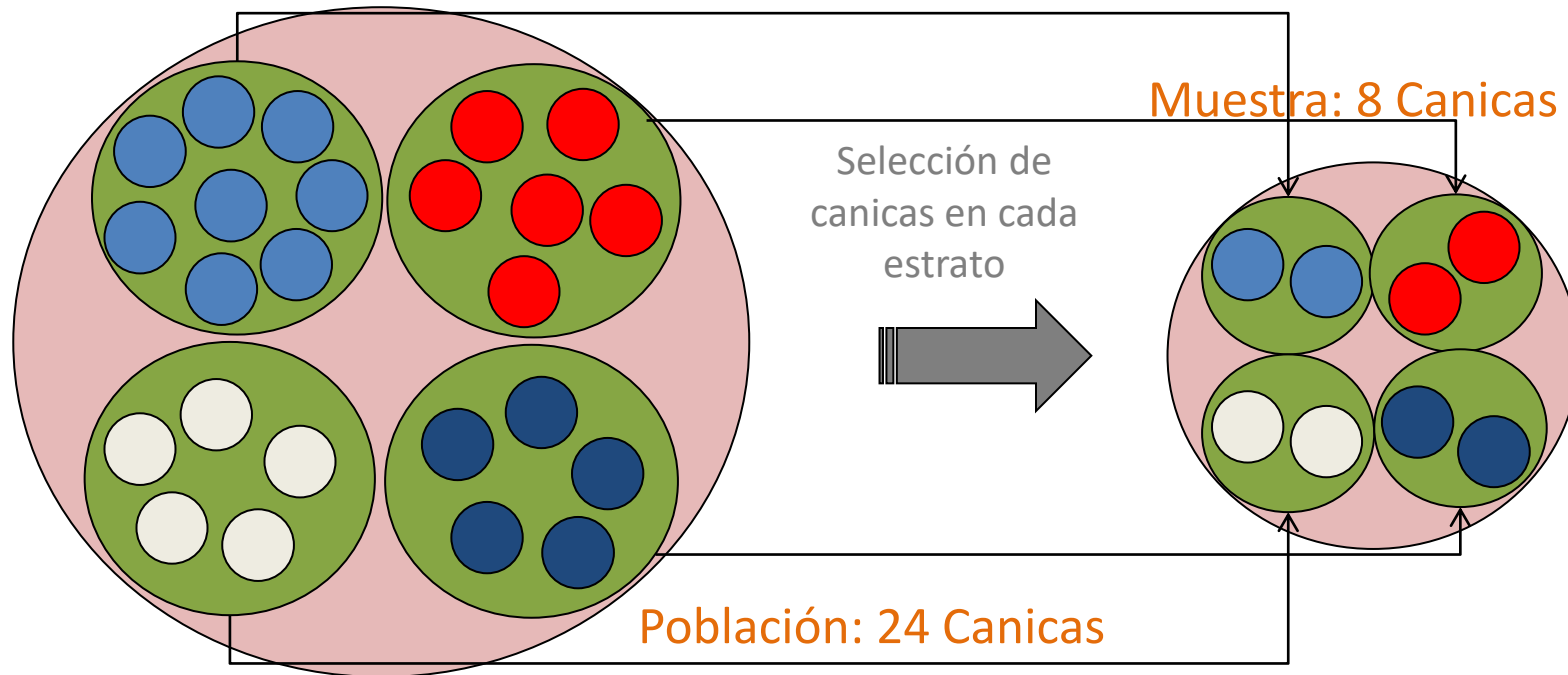
Se selecciona un punto aleatorio de inicio y posteriormente se elige cada k-ésimo miembro de la población.



Este procedimiento de selección es muy útil y fácil de aplicar e implica el seleccionar dentro de una población N a un número n de elementos a partir de un intervalo $K=N/n$.

Muestreo Estratificado

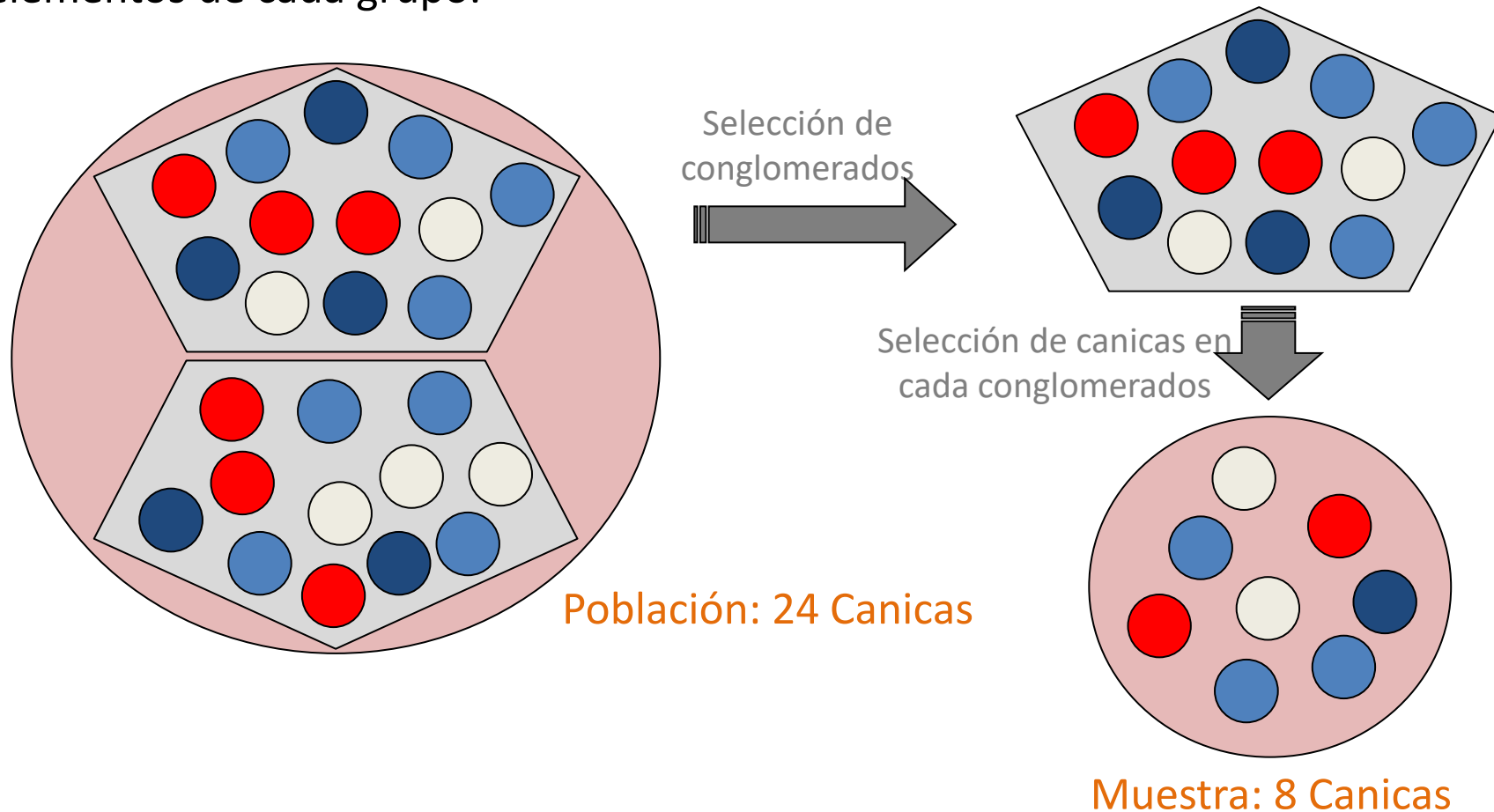
Una población se divide en subgrupos, denominados estratos y se selecciona al azar una muestra de cada estrato.



Se buscan grupos homogéneos al interior y heterogéneos entre grupos.

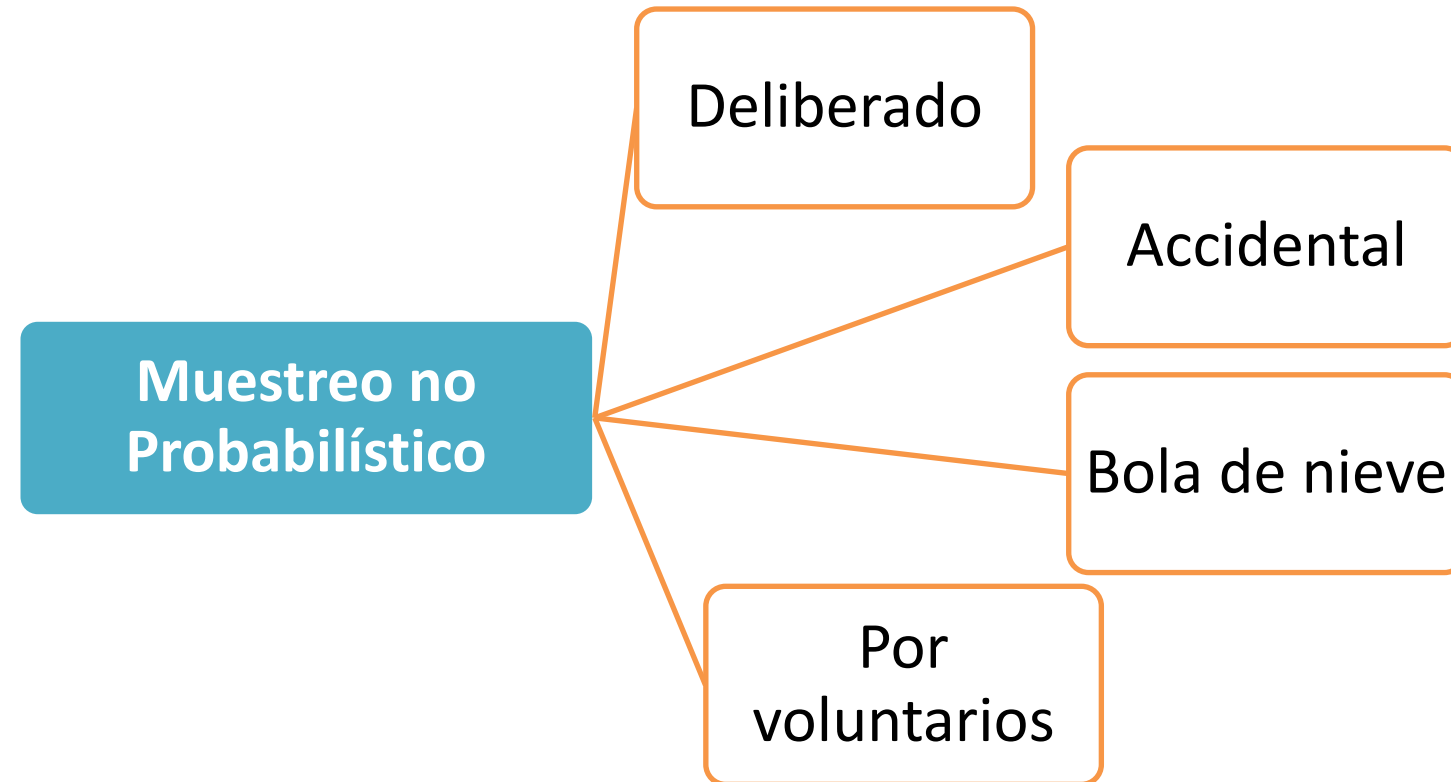
Muestreo por Conglomerado

La población se divide en conglomerados a partir de los límites naturales geográficos o de otra clase. A continuación se seleccionan los conglomerados al azar y se toma una muestra de forma aleatoria con elementos de cada grupo.



Muestreo No-Probabilístico

Este tipo de muestreo no sirve para realizar generalizaciones (estimaciones inferenciales sobre la población), pues no se tiene certeza de que la muestra sea representativa, ya que no todos los sujetos de la población tienen la misma probabilidad de ser elegidos.



Muestreo Deliberado

Muestreo no probabilístico en donde los criterios de selección se basan en el criterio personal del investigador a cerca de la representatividad de la población que se esta estudiando.

Por ejemplo, la mayoría de investigaciones de mercados de prueba de productos que se llevan a cabo en centros comerciales, son típicos del muestreo deliberado.

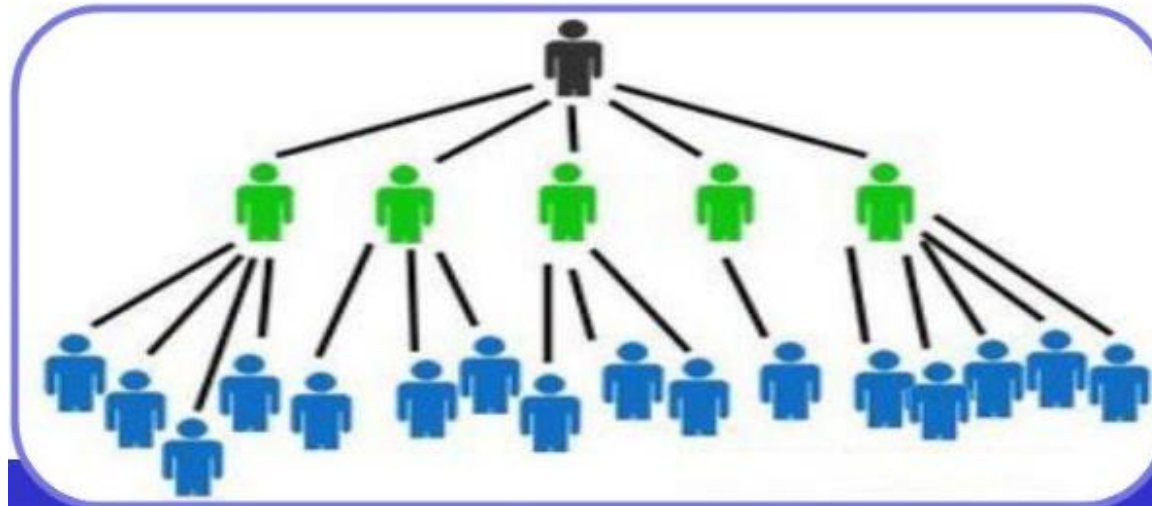
Muestreo Accidental

Muestreo no probabilístico donde el investigador elige a aquellos individuos que están a mano.

Por ejemplo, un periodista que va por la calle preguntando a las personas que salen a su paso, sin atender ningún criterio especial de elección.

Muestreo Bola de Nieve

Sistema de muestreo de conveniencia que consiste en la capacidad del investigador para encontrar un conjunto de individuos con las características deseadas. Estos individuos se utilizarán como informadores para identificar a otras personas con las características deseadas, y éstas a otras, y así sucesivamente.



Muestreo por Voluntarios

Es una muestra que consiste en unidades de la población que eligen voluntariamente responder o participar en la investigación.

Suele usarse en caso de intervenciones en que por sus características, molestas, peligrosas o dolorosas requiere de una participación voluntaria.

Ejercicio

De acuerdo a los tipos de muestreo probabilísticos determine cuál sería el tipo de muestreo que podría utilizarse en cada situación:

Se tiene un embarque de 1000 cajas con 50 focos en cada una. Existe el interés de estimar el número de focos defectuosos en cada caja.



Ejercicio

De acuerdo a los tipos de muestreo probabilísticos determine cuál sería el tipo de muestreo que podría utilizarse en cada situación:

En un vuelo de un jumbo se debe entrevistar a algunos pasajeros, sobre el servicio que presta la aerolínea, (a cada viajero se le asigna un puesto numerado).



Ejercicio

De acuerdo a los tipos de muestreo probabilísticos determine cuál sería el tipo de muestreo que podría utilizarse en cada situación:

Medir el nivel de satisfacción de los usuarios sobre los servicios públicos ofrecidos por EMCALI en la ciudad de Cali.



Ejercicio

De acuerdo a los tipos de muestreo probabilísticos determine cuál sería el tipo de muestreo que podría utilizarse en cada situación:

Una empresa pequeña que cuenta con 20 empleados desea medir el clima organizacional en su personal.

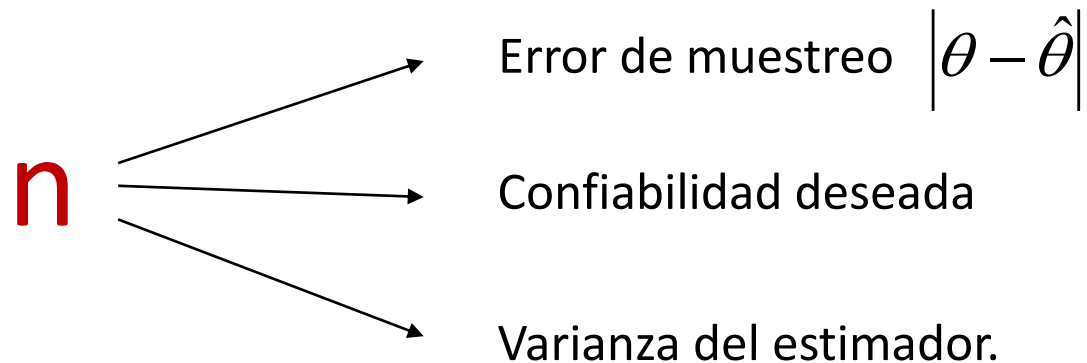


Tamaño de la Muestra

Tamaño de la Muestra

Una de las decisiones más importantes que deben tomarse en un plan de muestreo es la determinación del tamaño optimo de muestra ***n***

De que depende ***n***?



*“El tamaño de la muestra es proporcional a la varianza del estimador,
en ningún caso proporcional al tamaño de la población”*

Cálculo del tamaño de muestra (n)

Para estimar la
Media Poblacional

$$n = \frac{z^2 \sigma^2}{\varepsilon^2}$$

- n : Es el tamaño de la muestra.
- Z : Es el percentil obtenido bajo el supuesto de normalidad para un nivel de confianza establecido
- σ^2 : La varianza de la población
- ε : El margen de error o error de muestreo máximo permitido
- π : La proporción de la característica de interés de la población

Para estimar la
Proporción de la Población

$$n = \frac{z^2 \pi (1 - \pi)}{\varepsilon^2}$$

Fracción de Muestreo:

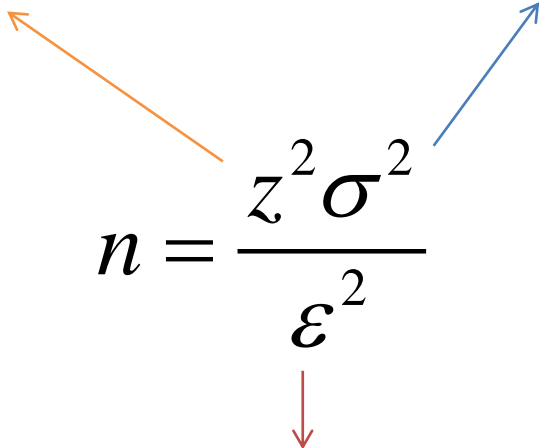
Si se conoce el tamaño de la población **N** y $\frac{n}{N} > 0.05$, se tendrá que:

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Sensibilidad del tamaño de muestra...

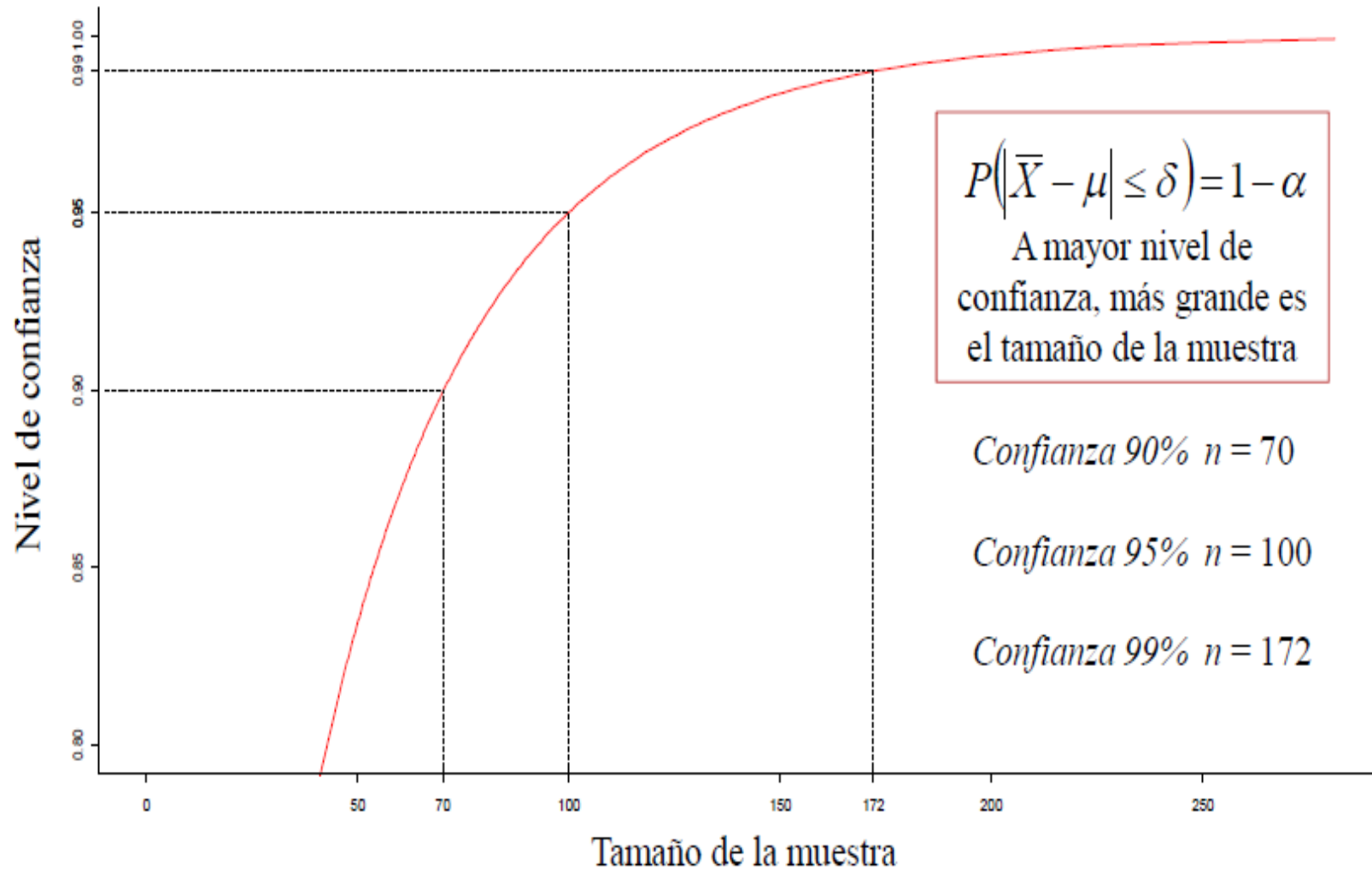
A mayor nivel de confianza, más grande es el tamaño de la muestra.

Entre más heterogénea sea la población, el tamaño de muestra requerido será mayor.

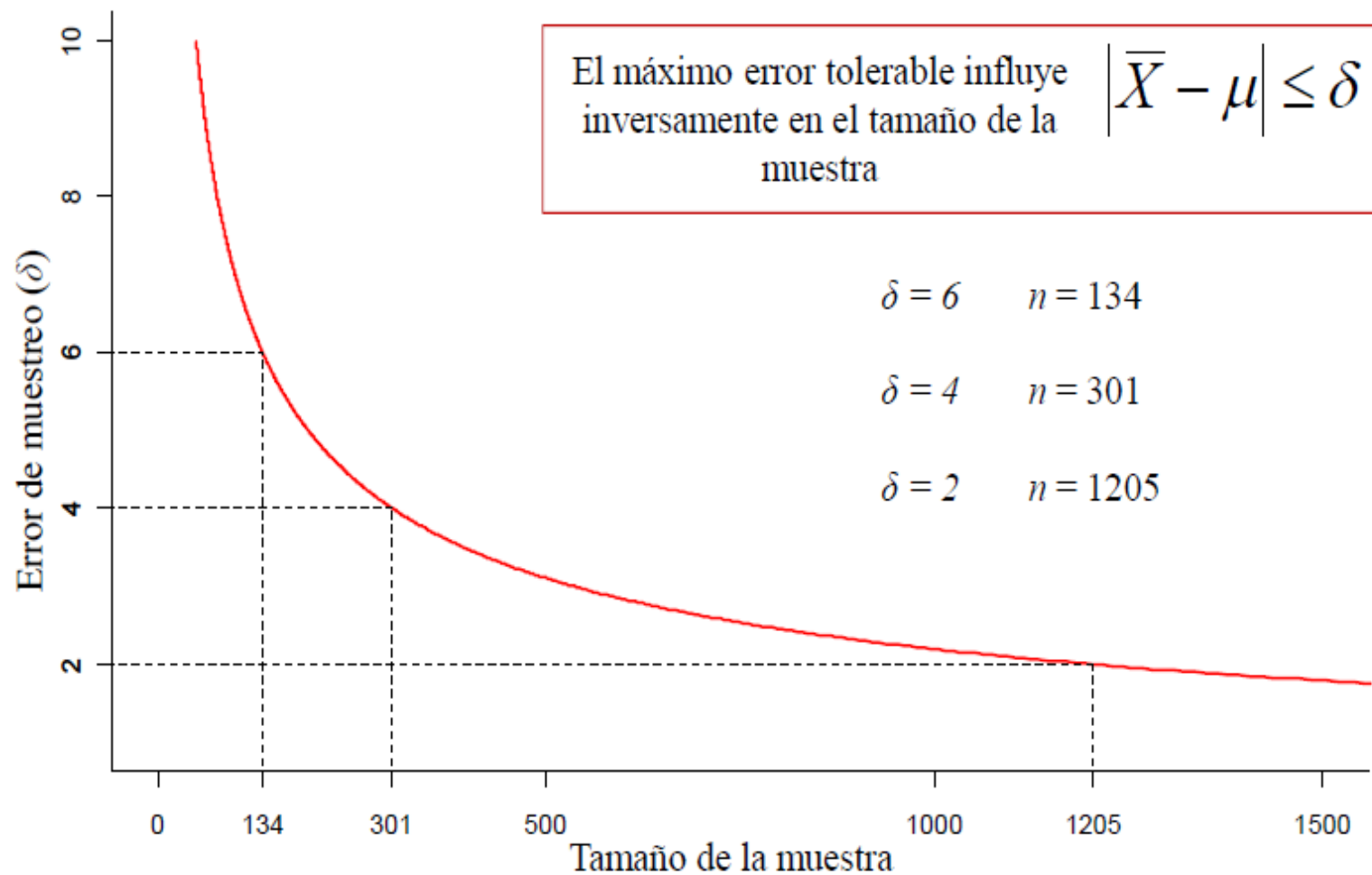
$$n = \frac{z^2 \sigma^2}{\varepsilon^2}$$


El máximo error tolerable influye inversamente en el tamaño de la muestra.

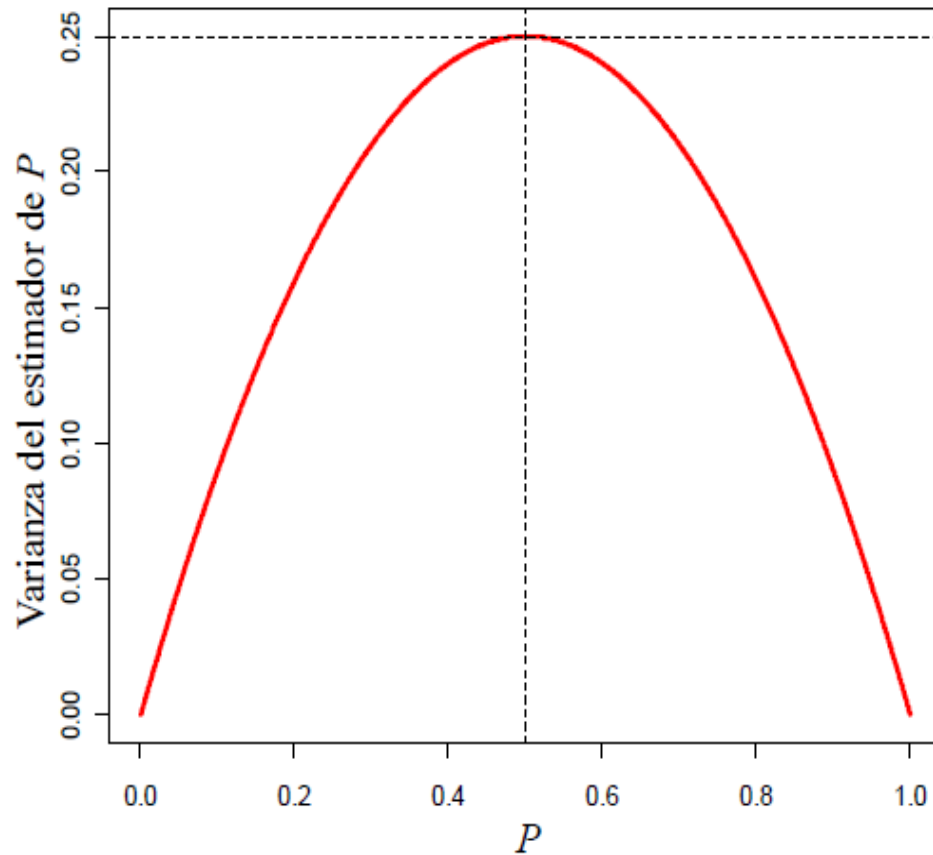
n vs Nivel de Confianza



n vs Error de Muestreo



Comportamiento de la Varianza para el Estimador de P



La varianza del estimador de P es mas alta para valores cercanos a 0.5.

La varianza del estimador de P es simétrica con respecto al valor de 0.5.

Ejemplo

En una comunidad integrada por 20.000 casas se desea estimar el consumo promedio mensual de gas por vivienda con un nivel de confianza del 95% y un error de muestreo de 6.95m³. Estudios anteriores para dicha comunidad permiten suponer que para el consumo mensual de gas en dicha comunidad $S^2=1255$.

¿Qué tamaño de muestra será necesario en este caso para llevar a cabo la correspondiente estimación?

$$n = \frac{Z^2 \cdot S^2}{\varepsilon^2} = \frac{(1.96)^2 \cdot 1255}{(6.95)^2} = 99.81 \approx 100$$

Como $\frac{n}{N} = \frac{99.81}{20000} = 0.00049905$, no es necesario aplicar la fracción de muestreo y el tamaño de muestra que garantiza el nivel de confianza y margen de error establecidos es $n=100$ casas.

$$N = 20.000$$

$$\varepsilon = 6.95$$

$$s^2 = 1255$$

$$NC = 95\%$$

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$\alpha/2 = 0.025$$

$$Z = 1.96$$

Si se aplicará la fracción de muestreo:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{99.81}{1 + \frac{99.81}{20.000}} = 99.31 \approx 100$$

Ejemplo

Se quiere estimar el porcentaje de estudiantes que quedan en embarazo durante su periodo de estudios en la Universidad, para esto se requiere tomar una muestra de estudiantes para indagar esto. Se sabe por estudios anteriores que este porcentaje ha sido cercano al 20%.

Estimar una muestra para cada facultad con una confianza del 95% y un error del 10%.

- Ciencias Económicas y Administrativas =1200
- Humanidades y ciencias Sociales= 1050
- Ciencias de la Salud=350
- Ingeniería=400

Ejemplo

Se quiere estimar el porcentaje de estudiantes que quedan en embarazo durante su periodo de estudios en la Universidad, para esto se requiere tomar una muestra de estudiantes para indagar esto, se sabe que el total de mujeres de la Universidad es de 3000. Por estudios anteriores que este porcentaje ha sido cercano al 50%. Estimar una muestra con una confianza del 95% y un error del 5%.

$$\begin{aligned} N &= 3000 \\ \varepsilon &= 5\% = 0.05 \\ \pi &= 0.50 \\ NC &= 95\% \end{aligned}$$

$$\begin{aligned} 1 - \alpha &= 0.95 \\ \alpha &= 0.05 \\ \alpha/2 &= 0.025 \\ Z &= 1.96 \end{aligned}$$

$$n = \frac{z^2 \pi (1 - \pi)}{\varepsilon^2} = \frac{1.96^2 * 0.5 * 0.5}{0.05^2} = 384.1 = 385$$

Como $\frac{n}{N} = \frac{385}{3000} > 0.128 \approx 12.8\%$, se debe aplicar la fracción de muestreo

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{385}{1 + \frac{385}{3000}} = 341.2 \approx 342$$

Se deben entrevistar 342 estudiantes.

Ejercicio

Las encuestas anteriores revelan que 30% de los turistas que van a Las Vegas a jugar durante el fin de semana gasta más de \$1.000 cada uno. La gerencia desea actualizar este porcentaje.

- a. El nuevo estudio utilizará el nivel de confianza de 90%. El estimador estará a menos de 1% de la proporción de la población. ¿Cuál es el tamaño necesario de la muestra?
- b. La gerencia indicó que el tamaño de la muestra determinado es demasiado grande. ¿Qué se puede hacer para reducir la muestra? Con base en su sugerencia, vuelva a calcular el tamaño de la muestra.

Ejercicio

Las encuestas anteriores revelan que 30% de los turistas que van a Las Vegas a jugar durante el fin de semana gasta más de \$1.000 cada uno. La gerencia desea actualizar este porcentaje.

$$\pi = 0.30$$

- a. El nuevo estudio utilizará el nivel de confianza de 90%. El estimador estará a menos de 1% de la proporción de la población. ¿Cuál es el tamaño necesario de la muestra?

$$n = \frac{z^2 \pi (1 - \pi)}{\varepsilon^2} = \frac{1.64^2 * 0.3 * 0.7}{0.01^2} = 5682$$

$$\varepsilon = 1\% = 0.01$$

$$NC = 90\%$$

$$1 - \alpha = 0.90$$

$$\alpha = 0.10$$

$$\alpha/2 = 0.05$$

$$Z = 1.64$$

- b. La gerencia indicó que el tamaño de la muestra determinado es demasiado grande. ¿Qué se puede hacer para reducir la muestra? Con base en su sugerencia, vuelva a calcular el tamaño de la muestra.

Para disminuir el tamaño de muestra se pueden realizar cualquiera de las siguientes acciones:

- Disminuir el nivel de confianza.
- Aumentar el margen de error.

Dado que el N.C. es del 90% y el $\varepsilon = 1\%$ siendo bastante exigente, se puede pensar entonces en aumentar el margen de error hasta el 5%. Miremos que pasaría:

$$n = \frac{z^2 \pi (1 - \pi)}{\varepsilon^2} = \frac{1.64^2 * 0.3 * 0.7}{0.05^2} = 227$$

Ejercicio

Si las directivas del programa de Gestión Empresarial desean hacer un seguimiento a las calificaciones de los estudiantes en los exámenes finales, sabiendo que la desviación estándar de estas calificaciones es de 1.8 y que la proporción de estudiantes que pierden los exámenes finales es de 0.43.

- a. Calcule el tamaño de muestra requerido para estimar el promedio de las calificaciones, con un nivel de confiabilidad del 98% y un error de muestreo de 0.2.
- b. Calcule el tamaño de muestra requerido para estimar la proporción de estudiantes que pierden los exámenes finales con, con un nivel de confiabilidad del 98%, un error de muestreo de 0.07.
- c. Para realizar el estudio conjunto sobre las calificaciones y el porcentaje de estudiantes que pierden los exámenes finales, qué tamaño de muestra utilizaría usted?

Ejercicio

Según información anterior, el índice de masa corporal promedio de los empleados es de 24,5 y la desviación típica es de 4,9.

¿Cual debe ser el tamaño de muestra apropiado si se desea estimar el índice de masa corporal en los nuevos empleados de la empresa, con una confianza del 95% y un margen de error de 0,8?

El valor de Z para obtener una confianza del 95% es 1,96

$$n = \frac{(1.96)^2 (4.9)^2}{0.8^2} = 144.11 \approx 144$$