UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

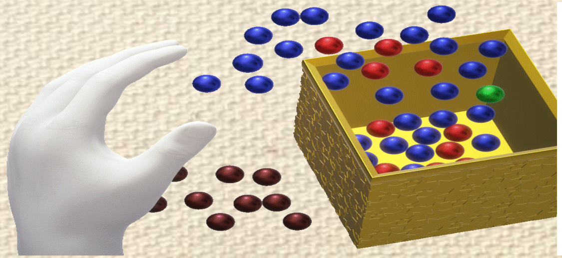
PROGRAMA DE POSGRADO EN CIENCIAS DE LA ADMINISTRACIÓN

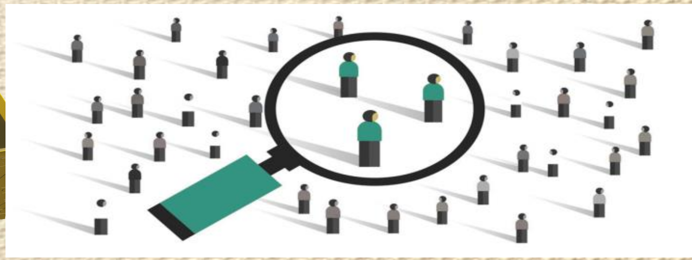
MAESTRÍA EN AUDITORÍA

Métodos Cuantitativos Aplicados a la Auditoría

Dra. María del Rosario Granados Sánchez

marzo – abril 2022





Introducción al muestreo

3.1 Concepto de muestreo. Muestreo estadístico y no estadístico.
3.2 Importancia del muestreo en auditoría, tanto en las pruebas de cumplimiento como en las pruebas sustantivas.
3.3 Boletines alusivos al muestreo en la normatividad del Instituto Mexicano de Contadores Públicos.
3.4 Diferentes tipos de muestreo estadístico.
3.5 Errores atribuibles al muestreo y no atribuibles al muestreo
3.6 Teorema central del límite
3.7 Niveles de confianza
3.8 Determinación de tamaños de muestra

Muestreo: Técnica basada en probabilidad y estadística, para seleccionar de un conjunto de elementos, únicamente una parte de ellos, de manera que los elementos seleccionados puedan proporcionar información de TODO el conjunto al cual pertenecen sin necesidad de evaluarlos uno por uno.

TÉCNICAS DE MUESTREO

- ➢ NO PROBABILÍSTICO
  - ✓ Muestreo por Conveniencia
  - ✓ Muestreo por Juicio
  - ✓ Muestreo por Cuota
  - ✓ Muestreo por Bola de Nieve
- ➢ PROBABILÍSTICO
  - ✓ Muestreo Aleatorio Simple
  - ✓ Muestreo Aleatorio Estatificado
  - ✓ Muestreo Probabilístico Sistemático
  - ✓ Muestreo Probabilístico por Conglomerado

Importancia del muestreo en auditoría:

- Aplicar las técnicas de muestreo estadístico a la auditoría, es una forma de evaluar una partida sin necesidad de aplicar procedimientos de auditoría a todos los integrantes de dicha partida; considerando un grado de certidumbre en nuestro juicio.

- La inferencia, que se usa para hacer conclusiones de un total, basados en una fracción de ese total, debe ser tratada sin abuso, pues el carácter de las revisiones que practica el Órgano de Fiscalización Superior, obliga a señalar a cabalidad cuáles y cuántas partidas o elementos auditados tienen deficiencias.
- Se abusa de la inferencia si proyectamos una observación hacia un todas las partidas que conforman el rubro revisado, ya sean cuentas, operaciones, registros; en función de una muestra en la cual se han detectado deficiencias. Es preferible, si así lo aprecia el auditor y el supervisor de la auditoría prescindir del muestreo y entrar a la revisión absoluta del rubro de interés.
- El muestreo es únicamente una herramienta, no existe la obligación o la restricción de su uso. Es válido para el auditor, trabajar una auditoría seleccionando partidas directamente por criterios de prioridad como denuncias, señalamientos históricos, datos de revisiones de cuenta pública y posteriormente trabajar una muestra estadística en la que ha sido extraída previamente la o las partidas de importancia trascendental. De cierto es que no solo se existe el muestreo estadístico. La selección de objetos, entes o personas realizado a juicio del interesado, errático o intencional, son ejemplos de muestreo no estadístico y el auditor puede hacer uso de ellos en el momento que sea necesario.

Las pruebas de cumplimiento se diseñan para comprobar que los procedimientos de control interno estaban en operación durante el periodo auditado

Las pruebas sustantivas se elaboran para llegar a una conclusión respecto al saldo de una cuenta, sin importar los controles internos sobre los flujos de las transacciones que se reflejan en el saldo.

Boletines alusivos al muestreo en la normatividad del Instituto Mexicano de Contadores Públicos

- **Boletín 6020: "El muestreo en la auditoria"** señala los principales aspectos relativos a la selección de partidas para ser revisadas en una auditoría de estados financieros, lo que se denomina muestreo, en el cual se requiere que todas las partidas que integran el universo sujeto a ser revisado tengan la misma oportunidad de ser seleccionadas, lo que dará como resultado la obtención de una muestra representativa del mismo.
- El objetivo de este boletín es establecer técnicas y lineamientos sobre el uso de sistemas de muestreo en la auditoría y el uso de otros medios de selección de partidas para reunir suficiente evidencia y evaluar sus resultados.
- **Boletín 6030: "Muestreo estadístico en auditoria"** El propósito es que sirva de orientación al auditor externo en la aplicación del muestreo estadístico como parte del proceso de una auditoría de estados financieros de una entidad, en lo que respecta a la planeación y determinación de muestras para propósito de su auditoría, así como para la evaluación de los resultados obtenidos.

¿Qué es una población?

¿Qué es una muestra?

¿Por qué seleccionamos una muestra cuando hacemos investigación?

Recabar datos → Inferencia estadística ↓ Población

Ej. Elecciones 2021, X partido quiere saber que proporción de votantes registrados probablemente le dé su voto

Muestra = 400
Darán su voto =160

¿Cuál es la proporción de votantes a favor del candidato?

$$\hat{p} = \frac{a\,favor}{tamaño\ de\ la\ muestra} \qquad \hat{p} = \frac{160}{400} = 0.40$$

¿Los resultados de una muestra son los resultados reales?

- Estimación
- No esperamos que el resultado se cumpla
- Solo contiene a una parte de la población, se espera algún error de muestreo

Pero…
Método adecuado de muestreo
Buenas estimaciones de los parámetros poblacionales

Población muestreada: *Aquella de la cual se extrae la muestra*

¿Cuál será la población muestreada si queremos conocer las preferencias de los votantes?

Marco muestral:

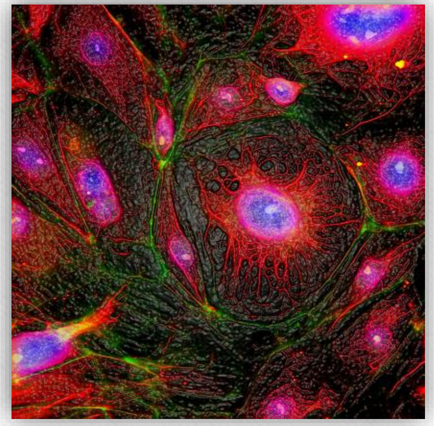
Lista de elementos de donde se obtendrá la muestra

Identificar físicamente a las unidades de análisis que conforman la población, así como enumerarlos.

¿Cómo es la población de estudio?

Infinita

Finita



Población → **Parámetros**

Media	μ
DS	σ
Proporción	p

Muestra → **Estimadores**

Media	$\bar{x}$
DS	s
Proporción	$\bar{p}$

Media$=\frac{\Sigma\, sueldos}{N}$ $\frac{129{,}500{,}000}{2500} = 51{,}800$

μ =51,800 USD

σ =4,000 USD

$p = \frac{1500}{2500} = 0.60$

¿Cuándo podemos trabajar con una población?

- **Población finita pequeña**

Ejemplo:
Supongamos que queremos conocer el perfil de los gerentes de una empresa.

N=2500
- **Sueldo anual (USD)**
- **Completó un curso de capacitación (1500)**

Conocemos toda la información
Tenemos los recursos necesarios para hacerlo:
- Información
- Tiempo
- Recursos humanos
- Recursos monetarios

MUESTRA

Selección de muestra en población finita

- **Muestreo Aleatorio Simple**

  Una MAS de tamaño "n" de una población finita de tamaño "N" es una muestra seleccionada de manera que cada posible muestra de tamaño "n" tenga la misma probabilidad de ser seleccionada.

¿Cómo lo hacemos?

- **Asignar número a cada sujeto de la población**
- **Auxiliarse de una tabla de números aleatorios**
- **Definir el número de dígitos con los cuales trabajaremos la tabla de números aleatorios**
- **Comenzar en cualquier lugar de la tabla y avanzar en la dirección que se defina**

Objetivo: encontrar a nuestros 30 gerentes que compondrán la muestra

63271	59986	71744	51102	15141	80714	58683	93108	13554	79945
88547	09896	95436	79115	08303	01041	20030	63754	08459	28364
55957	57243	83865	09911	19761	66535	40102	26646	60147	15702
46276	87453	44790	67122	45573	84358	21625	16999	13385	22782
55363	07449	34835	15290	76616	67191	12777	21861	68689	03263
69393	92785	49902	58447	42048	30378	87618	26933	40640	16281
13186	29431	88190	04588	38733	81290	89541	70290	40113	08243
17726	28652	56836	78351	47327	18518	92222	55201	27340	10493
36520	64465	05550	30157	82242	29520	69753	72602	23756	54935
81628	36100	39254	56835	37636	02421	98063	89641	64953	99337
84649	48968	75215	75498	49539	74240	03466	49292	36401	45525
63291	11618	12613	75055	43915	26488	41116	64531	56827	30825
70502	53225	03655	05915	37140	57051	48393	91322	25653	06543
06426	24771	59935	49801	11082	66762	94477	02494	88215	27191
20711	55609	29430	70165	45406	78484	31639	52009	18873	96927
41990	70538	77191	25860	55204	73417	83920	69468	74972	38712
72452	36618	76298	26678	89334	33938	95567	29380	75906	91807
37042	40318	57099	10528	09925	89773	41335	96244	29002	46453
53766	52875	15987	46962	67342	77592	57651	95508	80033	69828
90585	58955	53122	16025	84299	53310	67380	84249	25348	04332
32001	96293	37203	64516	51530	37069	40261	61374	05815	06714
62606	64324	46354	72157	67248	20135	49804	09226	64419	29457
10078	28073	85389	50324	14500	15562	64165	06125	71353	77669
91561	46145	24177	15294	10061	98124	75732	00815	83452	97355
13091	98112	53959	79607	52244	63303	10413	63839	74762	50289

6327, 8671, 7445…

Se descartan porque son mayores a 2500, por tanto, no corresponden a ningún gerente.

Ejercicio:

1. Tome una población finita con cinco elementos A, B, C, D y E. Se pueden seleccionar 10 muestras aleatorias simples de tamaño 2.
   a) Liste las 10 muestras empezando con AB, AC y así en lo sucesivo.
   b) Utilizando el muestreo aleatorio simple, ¿cuál es la probabilidad para cada muestra de tamaño 2 de ser seleccionada?
   c) Asuma que el número aleatorio 1 corresponde a A, el número 2 corresponde a B y así en lo sucesivo. Liste la muestra aleatoria de tamaño 2 que será seleccionada al usar los números aleatorios 8 0 5 7 5 3 2.
2. Suponga que una población finita tiene 350 elementos. A partir de los últimos tres dígitos de cada uno de los siguientes números aleatorios de cinco dígitos (por ejemplo: 601, 022, 448, . . .), determine los primeros cuatro elementos que se seleccionarán para una muestra aleatoria simple.

   98 601 73 022 83 448 02 147 34 229 27 553 84 147 93 289 14 209

3. *Fortune* publica datos sobre ventas, valor del activo, valor de mercado y utilidades por acción de las 500 corporaciones industriales más grandes de Estados Unidos (*Fortune* 500, 2006). Suponga que usted desea seleccionar una muestra aleatoria simple de 10 corporaciones de la lista *Fortune* 500. Use los tres últimos dígitos de la novena columna de la tabla 7.1, empezando con 554. Leyendo hacia abajo por esa columna, identifique los números de las 10 corporaciones que se tomarán para la muestra.
4. A continuación se presentan las 10 acciones más activas en la Bolsa de Nueva York del 6 de marzo de 2006 (*The Wall Street Journal* 7 de marzo de 2006).

AT&T	Lucent	Nortel	Qwest	BellSouth
Pfizer	Texas Instruments	General Electric	iShrMSJpn	LSI Logic

   Las autoridades bursátiles decidieron investigar las prácticas de negociación utilizando una muestra de tres de estas acciones.
   a) Comenzando con el primer dígito aleatorio de la sexta columna de la tabla 7.1, lea los números descendiendo por esa columna para seleccionar una muestra aleatoria simple de tres acciones para las autoridades.
   b) Con la información aportada en la nota y comentario 3, determine cuántas muestras aleatorias simples diferentes de tamaño 3 pueden seleccionarse de una lista de 10 acciones.

Combinaciones

- **TIPOS DE MUESTREO**

NO PROBÁBILISTICO

Métodos no probabilísticos, aun siendo conscientes de que no sirven para realizar generalizaciones, pues no se tiene certeza de que la muestra extraída sea representativa, ya que no todos los sujetos de la población tienen la misma probabilidad de se elegidos. En general se seleccionan a los sujetos siguiendo determinados criterios procurando que la muestra sea representativa.

Muestreo Intencional

Es cuando una persona selecciona la muestra procurando que sea representativa, dependiendo tal representatividad de su intención u opinión, con lo que **la evaluación es subjetiva**. En este proceso no se produce una selección aleatoria de las muestras limitándose el muestreo a unidades que parecen ser representativas de la población que se considera. Se obtiene información sobre esas unidades y con base en la misma se hacen estimaciones sobre las características de la población.

Aplicando Criterio.

El criterio de la persona que selecciona la muestra es importante, porque personas diferentes tiene criterios diferentes. No hay método objetivo por el que se prefiera un criterio de selección de muestra a otro.

Muestreo sin norma.

Es cuando se toma la muestra de cualquier manera por razones de comodidad. La representatividad de tal muestra sólo puede aspirar a ser medianamente satisfactoria en el caso de que la población sea homogénea.

Muestreo por cuotas o accidental:

Se asienta generalmente sobre la base de un buen conocimiento de los **estratos de la población más "representativos"** o "adecuados" para los fines de la investigación. Mantiene, por tanto, **semejanzas con el muestreo aleatorio estratificado, pero no tiene el carácter de aleatoriedad de aquél.**

Se fijan unas "cuotas" que consisten en un número de individuos que reúnen unas determinadas condiciones, por ejemplo: 30 personas de 25 a 40 años, de sexo femenino y residentes en la Ciudad de Monterrey. Una vez determinada la cuota se eligen los primeros que se encuentren que cumplan esas características. Este método se utiliza mucho en las encuestas de opinión.

Muestreo opinático o 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 preelectorales de zonas que en anteriores votaciones han marcado tendencias de voto.

Bola de nieve:

Localiza a algunos individuos, los cuales conducen a otros, y estos a otros, y así hasta conseguir una muestra suficiente. Este tipo se emplea muy frecuentemente cuando se hacen estudios con poblaciones "marginales", delincuentes, sectas, determinados tipos de enfermos, etc.

Muestreo casual o incidental:

Se trata de un proceso en el que **el investigador selecciona directa e intencionadamente los individuos de la población**. El caso más frecuente de este procedimiento el utilizar como muestra los individuos a los que se tiene fácil acceso (los profesores de universidad emplean con mucha frecuencia a sus propios alumnos). Un caso particular es el de los voluntarios.

- **Muestreo probabilístico**

Muestreo de una población infinita

- No es posible elaborar una lista de todos los elementos de la población

MUESTRA aleatoria

De tamaño “n” de una población infinita debe cumplir las siguientes condiciones:

a. Cada elemento seleccionado proviene de la misma población
b. Cada elemento es seleccionado de manera independiente

Ejemplo de los neumáticos

Proceso productivo no hay un límite en el número de elementos que se genera

Personas en un restaurante

Que sean comensales

Omitimos a los que van en busca del sanitario

No ser del mismo grupo de clientes

No generar sesgo

Cereales

Selección de muestras de tamaño 12

Llenado de cajas con un peso medio de 24oz

Determinar si funciona correctamente: sobrellenado o menor peso

¿Cómo seleccionamos la muestra?

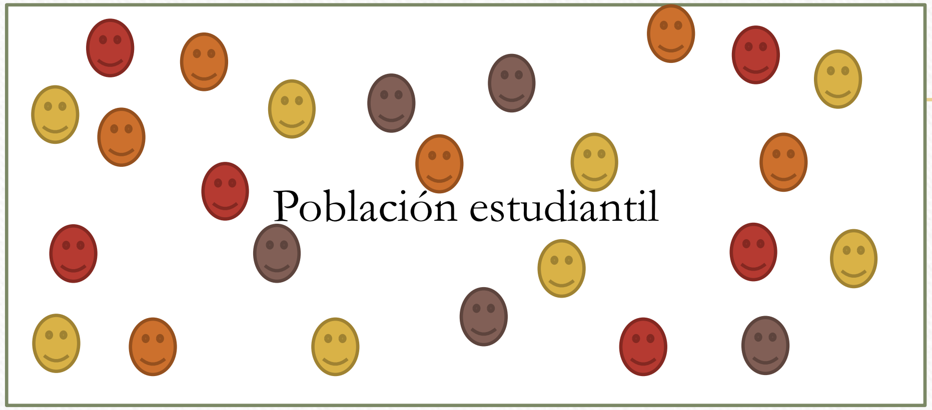
- Los elementos seleccionados vienen de la misma población
- Cada elemento es seleccionado de manera independiente

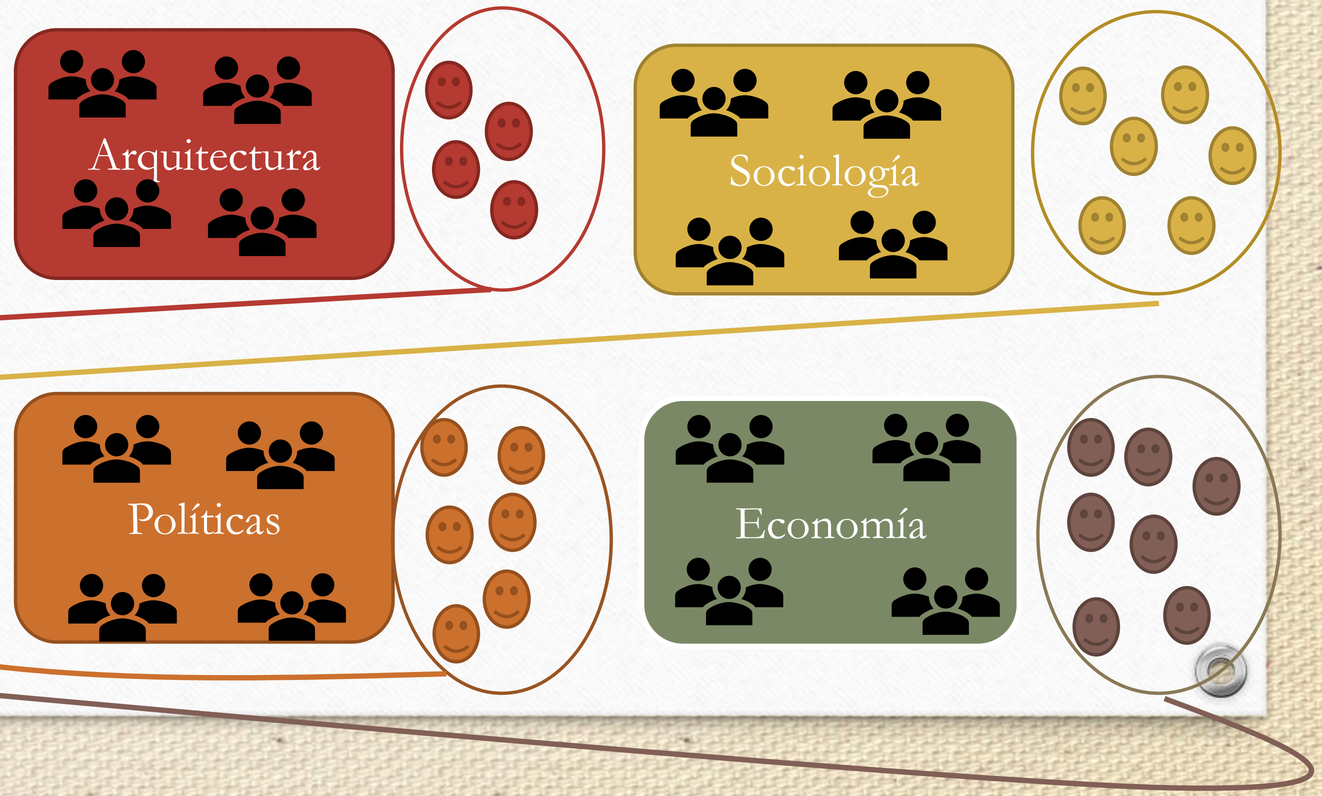
Muestreo Aleatorio Estratificado

Población se divide en subgrupos (estratos), y se selecciona una muestra de cada uno.

Segmentos homogéneos

Un individuo solo puede pertenecer a un estrato y luego a través de muestreo simple se genera la muestra de cada segmento





Muestra

Muestra proporcional

Muestra no proporcional

Número de elementos de cada estrato tenga la misma proporción que se encuentra en la población

Empresas: N=352
Estratos generados en función de los dividendos: 5
Muestra de tamaño n=50

Estrato	Dividendos	# Empresas	% del total	Cantidad muestreada
1	> 30%	8		
2	20 al 30%	35		
3	10 al 20%	189		
4	0 al 10%	115		
5	< 0 (déficit)	5		
Total		352		

Muestra no proporcional

La cantidad de elementos estudiada es desproporcionada respecto de su número en la población.

Se ponderan los resultados de la muestra de acuerdo con la proporción del estrato respecto a la población total.

Estrato	Dividendos	# Empresas	% del total	Cantidad muestreada
1	> 30%	8	2	
2	20 al 30%	35	10	
3	10 al 20%	189	54	
4	0 al 10%	115	33	
5	< 0 (déficit)	5	1	
Total		352		

/ 100 =

Los estratos que más se utilizan son:

- Religión
- Edad
- Nacionalidad
- Género
- Nivel educativo, etc…

Independientemente del procedimiento de muestreo, cada elemento de la población tiene probabilidad de ser seleccionado para la muestra

Fortalezas:

1. Refleja con más precisión las características de la población
2. Representación más equitativa
3. Se pueden observar relaciones entre dos grupos y diferenciar comportamientos dentro de la población

Debilidades:

1. Más caro, tardado y tedioso
2. Conocer bien la población para separar por estratos
3. Generar sesgo

Si la condición de estratos homogéneos internamente y heterogéneos entre si, se cumple.
Mejora la precisión de resultados al realizar un estudio sobre la muestra.

Muestreo por conglomerados (clusters)

Las unidades de muestreo son grupos de unidades de estudio.

Elementos divididos en grupos separados (conglomerados o clusters)

Cada elemento pertenece solo a un conglomerado

MAS de los conglomerados

Los elementos del conglomerado forman la muestra

Cada conglomerado sea una representación de la población completa.

Hay mejores resultados con elementos no semejantes en los conglomerados

Buena estimación parámetros poblacionales

¿Cuándo utilizarlo?

No hay estratos bien definidos
Cuando se está muestreando dentro de ellos
Cuando la población es grande y está conformada por subconjuntos
No hay un marco muestral

1. Definir conglomerados
2. Se seleccionan a la azar algunos de ellos para estudiarlos
3. Se puede definir a través de un criterio geográfico
4. Podemos dividir el total de la población en provincias
5. Seleccionar los conglomerados a estudiar (MAS)
6. Aplicar encuestas o entrevistas a los miembros del conglomerado
7. Nuevo proceso de muestreo dentro del conglomerado

Muestreo Unietápico

Muestreo Bietápico

Adecuado cuando hay grupos similares entre si, por lo que no hay diferencia entre estudiar individuos de un grupo u otro.

Ventajas:

Operativa, reduces costos de muestreo en áreas geográficas grandes.

Desventajas:

Riesgo de que los conglomerados no sean realmente homogéneos.

Ejemplo:

Queremos investigar el rendimiento académico de estudiantes de primaria en México

14.3 Millones de estudiantes de primaria

1. Población de estudiantes de primaria en México

2. Dividimos en ciudades, regiones, estados…

3. Seleccionamos los lugares a encuestar (M.A.S) (sorteo)

4. Sorteo

6, 9, 15, 31, 23,
22, 1

5. Encuestamos a los estudiantes de los conglomerados o elegimos un número al azar de cada conglomerado

MUESTRA

Error de muestreo

- Diferencia entre el estadístico de una muestra y el parámetro de la población correspondiente.

- El número de habitaciones rentadas en Foxtrot Inn, en Tryon, Carolina del Norte en el mes de junio fue de 94.
- Determine la media de la población. $\mu = \frac{\Sigma X}{N} =$
- Seleccionar tres muestras aleatorias de cinco días.
- Calcule la media de cada muestra y compárela con la media poblacional.

$$\overline{X}_1 = \frac{\Sigma X}{n} = \frac{4 + 7 + 4 + 3 + 1}{5} = \frac{19}{5} = 3.80$$

- ¿Cuál es el error de muestreo en cada caso?

Errores atribuibles al muestreo y no atribuibles al muestreo

- El auditor debe considerar los resultados de la muestra, la naturaleza y causa de cualquier error identificado y su posible efecto en el objetivo de la prueba particular y en otras áreas de la auditoría.
- El error tolerable es el error máximo en el universo que el auditor estaría dispuesto a aceptar y a pesar de eso, concluir que el resultado del muestreo ha alcanzado su objetivo de auditoría.
- Para procedimientos sustantivos, el error tolerable será un monto menor que, o igual, al estimado preliminar del auditor de la materialidad usando para saldos de cuentas individuales que están siendo auditados. A menor grado de error tolerable, será mayor el tamaño de la muestra que requerirá el auditor.
- En los procedimientos de cumplimiento, el error tolerable es el porcentaje máximo de desviación de un procedimiento de control prescrito, que el auditor estaría dispuesto a aceptar sin alterar la confianza que tenía depositada en el control que está probando.
- Si el auditor espera la presencia de error, conocido como error esperado en el universo, normalmente tendrá que examinar una muestra mayor para concluir que el valor del universo está razonablemente presentado dentro del error tolerable estimado o que la confianza que se había depositado en un control importante está justificada. Las muestras de menor tamaño se justifican cuando se espera que el universo se encuentre libre de errores.

Al analizar los errores detectados en la muestra, el auditor debe determinar que cada partida sobre la que se tenga duda sobre su corrección será, de hecho, un error. Al diseñar la muestra, el auditor tendrá definidas aquellas condiciones que constituyen un error por medio de referencia a los objetivos de su auditoría.

En aquéllos en que no se pueda localizar la documentación de apoyo de partidas específicas de la muestra, el auditor tal vez pueda obtener la evidencia de auditoría apropiada a través de la aplicación de procedimientos alternativos relacionados con las partidas no probadas de una muestra.

Si el auditor no aplica o no puede aplicar procedimientos alternativos en relación con partidas no probadas de una muestra (cuentas por cobrar a clientes), debe considerar la partida como un error para los propósitos de su evaluación de la evidencia de auditoría proporcionada por la muestra de auditoría. El auditor también debe considerar los aspectos cualitativos de los errores. Éstos incluyen la naturaleza y la causa del error y el posible impacto del error en otras fases de auditoría. Por ejemplo, el grado de confianza que se planeó depositar en los procedimientos de control interno contable.

El auditor también debe considerar los aspectos cualitativos de los errores; éstos incluyen la naturaleza y la causa del error y el posible impacto del error en otras fases de auditoría. Asimismo, el auditor debe razonar el efecto de los errores, si provienen de violaciones al control interno respecto de la posible existencia de fraude.

Al evaluar los errores detectados, el auditor puede llegar a la conclusión de que muchos de ellos tienen una característica en común. Debe, entonces, llevar a cabo evaluaciones por separado basadas en las partidas examinadas por cada sub-universo.

Para procedimientos sustantivos, el auditor debe proyectar los errores monetarios encontrados en la muestra de un universo, y también considerar el efecto del error proyectado en el objetivo de la revisión en particular y en otras áreas de la auditoría.

Distribución muestral de la muestra

Distribución de probabilidad de todas las posibles medias de las muestras de un determinado tamaño muestral de la población.

Ingresos por hora de empleados de producción en Tartus Industries

Empleado	Ingresos por hora	Empleado	Ingresos por hora
Joe	$7	Jan	$7
Sam	7	Art	8
Sue	8	Ted	9
Bob	8		

¿Cuál es la media de la población?
¿Cuál es la distribución muestral de la media de muestras de tamaño 2?
¿Cuál es la media de la distribución muestral de la media?
¿Qué observaciones es posible hacer sobre la población y la distribución muestral de la media?

$$\mu = \frac{\Sigma X}{N} = \frac{\$7 + \$7 + \$8 + \$8 + \$7 + \$8 + \$9}{7} = \$7.71$$

Muestra	Empleados	Ingresos por hora	Suma	Media	Muestra	Empleados	Ingresos por hora	Suma	Media
1	Joe, Sam	$7, $7	$14	$7.00	12	Sue, Bob	$8, $8	$16	$8.00
2	Joe, Sue	7, 8	15	7.50	13	Sue, Jan	8, 7	15	7.50
3	Joe, Bob	7, 8	15	7.50	14	Sue, Art	8, 8	16	8.00
4	Joe, Jan	7, 7	14	7.00	15	Sue, Ted	8, 9	17	8.50
5	Joe, Art	7, 8	15	7.50	16	Bob, Jan	8, 7	15	7.50
6	Joe, Ted	7, 9	16	8.00	17	Bob, Art	8, 8	16	8.00
7	Sam, Sue	7, 8	15	7.50	18	Bob, Ted	8, 9	17	8.50
8	Sam, Bob	7, 8	15	7.50	19	Jan, Art	7, 8	15	7.50
9	Sam, Jan	7, 7	14	7.00	20	Jan, Ted	7, 9	16	8.00
10	Sam, Art	7, 8	15	7.50	21	Art, Ted	8, 9	17	8.50
11	Sam, Ted	7, 9	16	8.00					

Distribución muestral de la media con $n = 2$

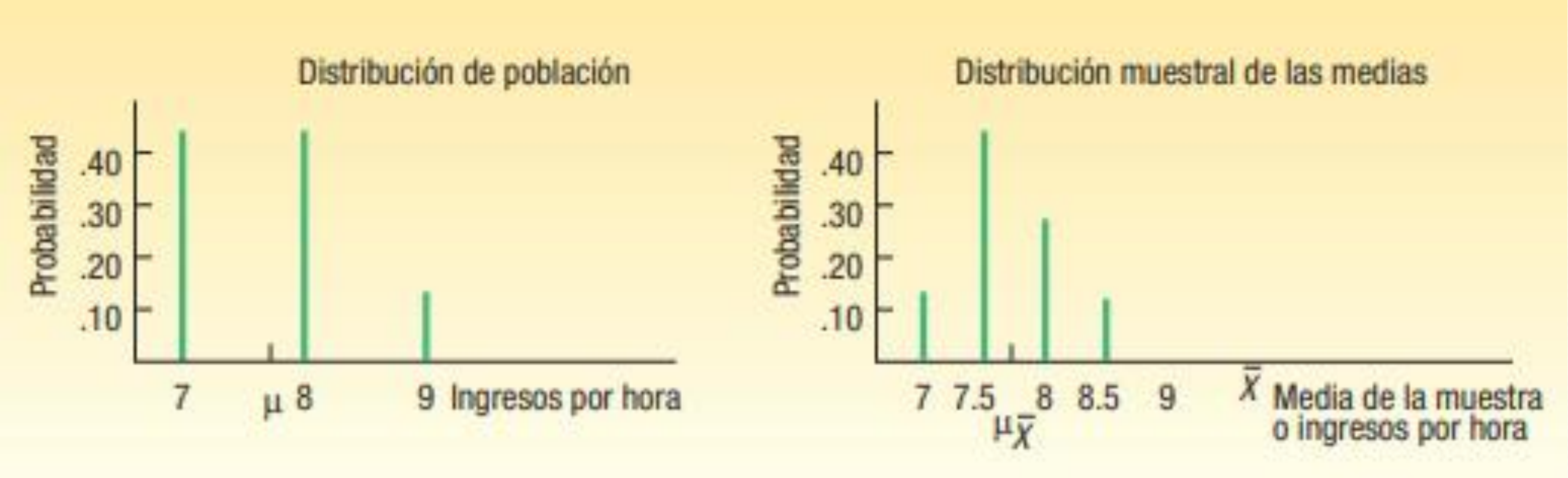
Media muestral	Número de medias	Probabilidad
$7.00	3	.1429
7.50	9	.4285
8.00	6	.2857
8.50	3	.1429
	21	1.0000

$$\mu_{\overline{X}} = \frac{\text{Suma de todas las medias muestrales}}{\text{Total de muestras}} = \frac{\$162}{21} = \$7.71$$

a) La media de la distribución muestral de la media ($7.71) es igual a la media de la población: $\mu = \mu_{\bar{X}}$.

b) La dispersión de la distribución muestral de las medias es menor que la dispersión de los valores de población. La media de las muestras varía de $7.00 a $8.50, mientras que los valores de población varían de $7.00 a $9.00. Observe que, conforme se incrementa el tamaño de la muestra, se reduce la dispersión de la distribución muestral de las medias.

c) La forma de la distribución muestral de la media y la forma de la distribución de frecuencias de los valores de población son diferentes. La distribución muestral de las medias tiende a adoptar más forma de campana y a aproximarse a la distribución de probabilidad normal.



Distribución de muestreo para medias muestrales

- La distribución de probabilidad de X se llama distribución muestral de la media.

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\boxed{E(\bar{X}) =} E\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} E(x_i)\right) = \frac{1}{n}\left(\sum_{i=1}^{n} \mu\right) = \frac{1}{n}(n\mu) = \boxed{\mu}$$

Tiene una distribución normal con media X. Cuando un estimador, como en este caso, tiene un valor esperado igual al parámetro que se desea estimar, se dice que es **un estimador insesgado**.

$$Var(x) = \frac{\sigma^2}{n}$$

$$z = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \approx N(0,1)$$

Existen tres propiedades importantes a considerar que son la base del teorema central de límite y que son las siguientes:

1.– $E(\bar{x}) = \mu$

2.- El error estándar para poblaciones infinitas es:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Ejemplo:

Calcular el **error estándar de una distribución muestral** de medias obtenida mediante la extracción de muestras de 500 personas, que provienen de una población que tiene una desviación de 25.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

¿Cuál es la desviación típica de una variable en una muestra de 350 personas que tiene un error de 12?

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma = \sigma_{\bar{x}}\left(\sqrt{n}\right)$$

Se desea saber el tamaño de la muestra que se debe de tener, cuando se sabe que el error es de 3 unidades y la desviación es de 15.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \sqrt{n} = \frac{\sigma}{\sigma_{\bar{x}}} \qquad n = \left(\frac{\sigma}{\sigma_{\bar{x}}}\right)^2$$

El error estándar para poblaciones finitas es:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Calcular el error estándar de una población de 300 personas, en relación a su estatura, cuando se extrae una muestra aleatoria de 145 personas cuya desviación estándar es de 8 cm.

3.- Sea cual fuera la forma de la distribución de la población, la forma de la distribución de muestreo de la media cuando n tiende a infinito siempre tenderá a la forma de la distribución normal.

Teorema del límite central

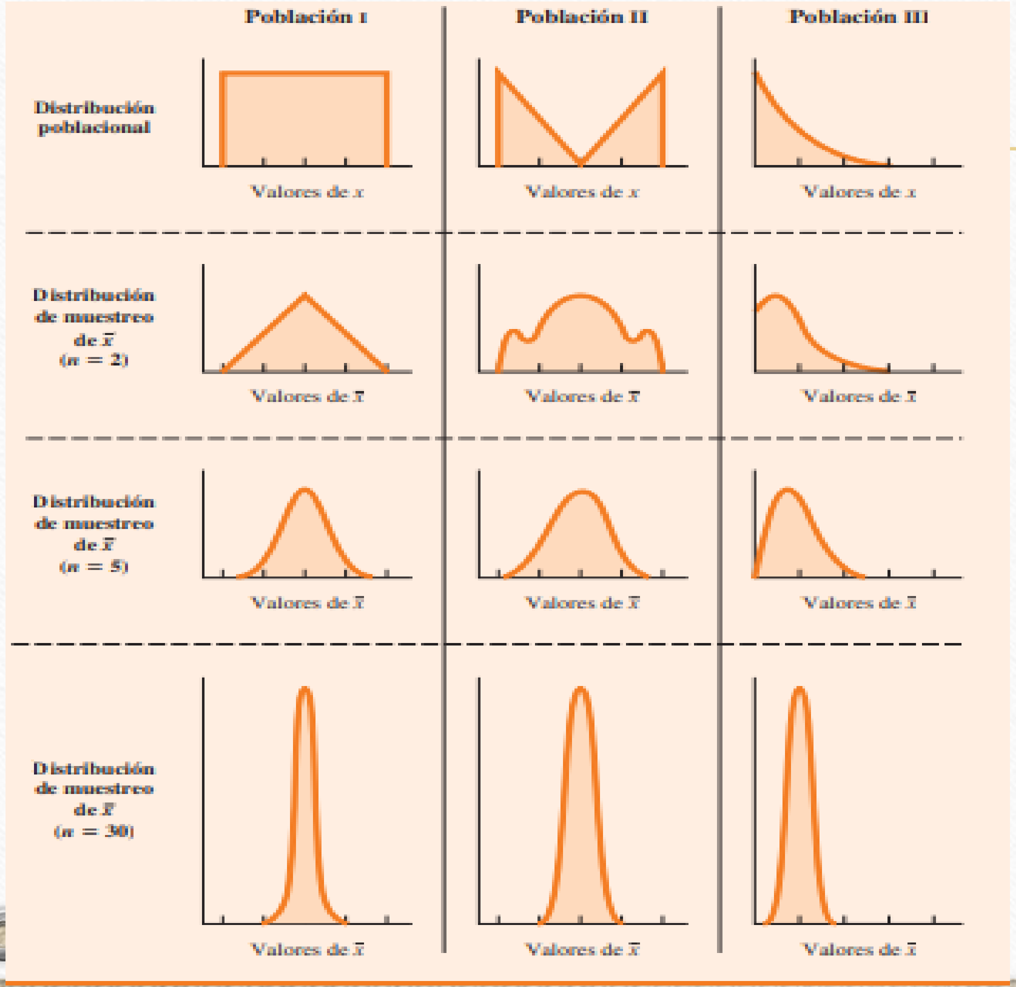
Forma de la distribución de muestreo de $(\bar{x})$

a. Población tiene una distribución normal

b. Población no tiene una distribución normal

Teorema del límite central.

Cuando se seleccionan muestras aleatorias de tamaño "n" de una población, la distribución de muestreo de la media muestral (x) puede aproximarse mediante una distribución normal a medida que el tamaño de la muestra se hace grande.



Ninguna población está distribuida normalmente

1. Distribución uniforme
2. Los valores se encuentran en las colas
3. Distribución exponencial: sesgada a la derecha

Cada distribución de muestreo tiene una forma diferente a la distribución poblacional.

Las distribuciones de muestreo empiezan a parecerse a la forma de una distribución normal

La forma de cada una de las tres distribuciones de muestreo es aproximadamente normal

Conforme n tiende a infinito es la distribución normal estándar n(z; 0, 1)

Si X es una variable aleatoria con media μ y desviación estándar σ, la distribución de muestreo de la media de la muestra $\bar{X}$ debe ser aproximadamente normal con media μ y desviación estándar sin importar la forma de la distribución original de la variable aleatoria X, siempre y cuando el tamaño de la muestra sea grande (n).

¿Qué tan grande debe ser el tamaño de la muestra antes de aplicar el teorema del límite central?

a. La distribución de muestreo de la media muestral se puede aproximar mediante una distribución normal siempre que la muestra sea de tamaño 30 o mayor
b. Cuando la población es muy sesgada o hay observaciones atípicas, pueden requerirse muestras de tamaño 50

Aplicación de la distribución de muestreo de $\bar{x}$

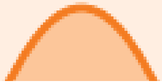
Siempre que seleccionamos una muestra aleatoria simple no podemos esperar que la media muestral sea exactamente igual a la media poblacional.

La distribución de muestreo de la x estriba en que se puede usar para proporcionar información probabilística acerca de la diferencia entre la media muestral y la media poblacional.

Suponga que el CEO de la empresa piensa que la media muestral será una estimación aceptable de la media poblacional, si la primera está en un margen de $500 de la segunda. .

En nuestro histograma algunas de las muestras (x) difieren en más de $2,000 de la σ.

¿Cuál es la probabilidad de que la media muestral obtenida usando una m.a.s de 30 gerentes se encuentre en un margen de $500 de la media poblacional?



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

1. Media poblacional=51,800

ál es la probabilidad de que la media
ral esté entre 51,300 y 52,300?

ongamos que la distribución de muestreo
istribuida normalmente

culamos el valor de Z en el extremo
or del intervalo (52,300)

$$= \frac{52300 - 51800}{730.3} = 0.68$$

culamos el valor de Z en el extremo inferior
ervalo (51,300)

$$= \frac{51300 - 51800}{730.3} = -\,0.68$$

- En un partida contable se sabe que la cantidad, en promedio, de estados financieros erróneos es de 4 con una desviación de 2, en una muestra de 30 días. Calcular la probabilidad de tener menos de 3 estados financieros con errores.

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Ubicar el valor Z en tablas:

Tamaño de muestra:

Población Finita

$$n = \frac{N * Z_{\alpha}^{2} * p * q}{e^{2} * (N-1) + Z_{\alpha}^{2} * p * q}$$

Población Infinita

$$n = \frac{Z_{\alpha}^{2} * p * q}{e^{2}}$$

n= tamaño de muestra
N: Tamaño de la población
Z: Parámetro estadístico que depende del nivel de confianza
e: Error de estimación máximo aceptado
p: Éxito
q: (1-p) fracaso

Z del 99% asociado a un α del 2.58%
Z del 95% asociado a un α del 1.96%
Z del 90% asociado a un α del 1.645%

"e" entre 1 y 10%

¿Cómo sabemos que "p" elegir?

¿Qué pasa si no la conocemos?

Ejercicio:
Calcular el tamaño de la muestra para una población de 32 estados de la República Mexicana para estudiar el rendimiento académico de estudiantes de primaria.
Nivel de confianza del 95%
Margen de error del 5%
Se desconoce la probabilidad del evento

N: 32; Z= 1.96; p= 50%; q= 50%; e= 5%

$$n = \frac{N * Z_{\alpha}^{2} * p * q}{e^{2} * (N-1) + Z_{\alpha}^{2} * p * q} \qquad n = \frac{32 * 1.96^{2} * 0.50 * 0.50}{0.05^{2} * (32-1) + 1.96^{2} * 0.50 * 0.50}$$

n= 29.61

384 ciudades

961 ciudades

¿Qué observan?

Entre más grande es la población, el tamaño de la muestra es relativamente pequeña

1. **Se consideraron el 93% de los estados**
2. **50% de las ciudades**
3. **28% de las ciudades**

- ## Tamaño de muestra para estimar la media de la población.
- Si empleamos el muestreo aleatorio simple, entonces:
- Partir de dos supuestos: en primer lugar el **nivel de confianza** al que queremos trabajar; en segundo lugar, cual es el **error máximo** que estamos dispuestos a admitir en nuestra estimación.

$$n_\infty = \frac{z_{\frac{\alpha}{2}}^2 \sigma^2}{e^2}$$

donde:

$z_{\frac{\alpha}{2}}$: z es el nivel de confianza elegido

σ^2 : varianza de la población

e: error máximo

Verificar que se cumpla:

$$N > n_\infty(n_\infty - 1)$$

Si esta condición se cumple el proceso termina aquí, y ese es el tamaño adecuado que debemos muestrear.

Si esto no se cumple, entonces el tamaño de la muestra se obtiene con la siguiente expresión:

$$n = \frac{n_\infty}{1 + \frac{n_\infty}{N}}$$

Una Compañía desea conocer la opinión de los empleados con respecto a vestir casual el viernes dentro del corporativo. La Compañía tiene un total de **10 000 empleados**, razón por la cual no puede entrevistar a todo el personal, con un **nivel de confianza del 95%,** un **error de 0.15** y una **varianza de 8.5**. Para tener una posición al respecto se obtendrá una muestra para tomar una decisión. **¿De qué tamaño debe de ser la muestra?**

$$n_{\infty} = \frac{z_{\frac{\alpha}{2}}^{2}\sigma^{2}}{e^{2}}$$

Se debe de cumplir con la siguiente condición:

$$\mathrm{N} > n_{\infty}(n_{\infty} - 1)$$

¿Se cumple la condición?

$$n = \frac{n_{\infty}}{1 + \frac{n_{\infty}}{N}}$$

Redondear siempre hacia arriba.

Distribuciones de muestreo para proporciones.

Es una distribución de tipo probabilístico que indica la medida de la probabilidad de que se presenten las proporciones de todas las muestras posibles del mismo tamaño de una población dada.

$$P = \mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i + \mu)^2}{N}}$$

P = proporción de la población
µ = Media de la población
x = número de éxitos
N = tamaño de la población
σ = desviación estándar de la población.

Media de la distribución de muestreo de la proporción: $P'' = E(\overline{P}) = P = \mu$

El error estándar para una población infinita: $\sigma_{\overline{P}} = \sqrt{\frac{P(1-P)}{n}}$

Calcular el error estándar de una proporción muestral si se tiene una muestra de 400 de una población cuya proporción es del 35%.

De que tamaño debe se ser la muestra cuando se sabe que la proporción es del 20% y el error típico es de 0.0238.

$$\sigma_{\overline{P}} = \sqrt{\frac{P(1-P)}{n}} \qquad n = \left(\frac{\sqrt{p(1-p)}}{\sigma_{\overline{P}}} \right)^2$$

El error estándar para una población finita:

$$\sigma_{\overline{P}} = \sqrt{\frac{P(1-P)}{n}} \sqrt{\frac{N-n}{N-1}}$$

Calcular el error estándar de una población de 250, de la cual se ha extraído una muestra de 90, se sabe que la proporción es del 15%.

$$\sigma_{\overline{P}} = \sqrt{\frac{P(1-P)}{n}} \sqrt{\frac{N-n}{N-1}}$$

Cuando n → infinito y P → 0.5 la distribución de muestreo de la proporción se aproxima a la forma de una distribución normal..

Para poblaciones infinitas:

$$z = \frac{(n\overline{P} \pm 0.5) - nP}{\sqrt{nP(1-P)}}$$

Para poblaciones finitas:

$$z = \frac{(n\overline{P} \pm 0.5) - nP}{\sqrt{nP(1-P)}\sqrt{\frac{N-n}{N-1}}}$$

Una compañía editorial, tiene 2500 distribuidores, se sabe que el 30% van a incrementar sus pedidos para el próximo período escolar. Calcular a probabilidad de que en una muestra de 250 distribuidores:

a.- una proporción del 35% o más incrementen sus pedidos:

$$z = \frac{((250)(0.35) - 0.5) - (250)(0.30)}{\sqrt{(250)(0.3)(1-0.3)}\sqrt{\frac{2{,}500 - 250}{2{,}500 - 1}}}$$

Buscar en tablas el valor de Z=

Por lo tanto $1 - 0.9505 = 0.0495$ entonces se tiene una probabilidad del 4.95%

- **Tamaño de muestra para estimar la proporción de la población.**
- Tener en cuenta los mismos factores que en el caso de la media:

$$n = \frac{N z_{\frac{\alpha}{2}}^{2} P(1-P)}{(N-1)e^{2} + z_{\frac{\alpha}{2}}^{2} P(1-P)}$$

$z_{\frac{\alpha}{2}}$: z valor que corresponde al nivel de confianza elegido.

P: proporción de una categoría de la variable

e: error máximo

N: tamaño de la población

Para establecer una política con respecto a los fumadores se desea realizar una encuesta a los empleados de la compañía BBC. Como no se puede encuestar a todos los empleados se desea conocer el tamaño de la muestra para tomar una posición al respecto. Los datos con los que se cuenta son: el 35% de la población fuma, el nivel de confianza es del 95% y el error máximo considerado es del 0.025. El total de la población es 10, 000.

$$n = \frac{(10{,}000)(1.96)^{2}(0.35)(1-0.35)}{(10{,}000-1)(0.025)^{2} + (1.96)^{2}(0.35)(1-0.35)}$$

Muestreo Aleatorio Simple

Población finita: una muestra seleccionada de tal manera que cada muestra posible de tamaño n tiene la misma probabilidad de ser seleccionada.

Población infinita: Una muestra seleccionada de tal manera que cada elemento proviene de la misma población y los elementos sucesivos se seleccionan de manera independiente.

Tamaño de la Muestra.

Necesitamos:

1. Varianza (Mientras más grande sea mayor será el tamaño de la muestra.)
2. Nivel de confianza (fijado por el auditor de acuerdo a su experiencia)
3. Precisión de la estimación (margen de error o error de muestreo que el investigador fija)

$$E = Z\frac{\sigma}{\sqrt{n}}$$ Población infinita

$$E = Z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$ Población finita

Cálculo de n en poblaciones infinitas.

4. Diseño muestral

$$Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

$$E = \bar{x}-\mu$$

$$Error = \frac{Z\sigma}{\sqrt{n}}$$

$$\sqrt{n} = \frac{Z\sigma}{E}$$

$$n = \left(\frac{Z\sigma}{E}\right)^2$$ Variable

$$n = \frac{Z^2PQ}{E^2}$$ Proporción

Cálculo de n en poblaciones finitas.

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$

$$E = \frac{Z\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N}} \Rightarrow E^2 = \left(\frac{Z^2\sigma^2}{n}\right)\left(\frac{N-n}{N}\right)$$

$$n = \frac{Z^2 N\sigma^2}{NE^2 + Z^2\sigma^2}$$

$$n = \frac{\sigma^2}{\left(\frac{E}{Z}\right)^2 + \frac{\sigma^2}{N}}$$

Un auditor desea tener un **nivel de confianza del 95%,** para que la verdadera proporción de **error no exceda del 2%.** Si la **población es muy grande**, ¿qué tamaño tendrá la muestra que va a tomarse, si el auditor estima que la **proporción de error es del 5%?**

$$n = \frac{Z^2 PQ}{E^2}$$

El mantenimiento de cuentas por cobrar puede resultar demasiado costoso, si el promedio de compra por cuenta baja de cierto nivel. El gerente de un almacén por departamentos desea estimar el promedio de lo comprado mensualmente por los clientes que usan la cuenta de crédito, con un **error de $1,500** y una **probabilidad aproximada del 0.95**. ¿Cuántas cuentas deberá seleccionar, si sabe que la **desviación estándar es de $30 000**, la cual fue obtenida de los balances mensuales de las cuentas de crédito?

$$n = \left(\frac{Z\sigma}{E}\right)^2$$

¿Qué tamaño de la muestra es necesario, si se considera una **confianza del 90%** para la proporción de la población, y el **error es del 8%?**

$$n = \frac{Z^2 PQ}{E^2}$$

Tamaño de la Muestra cuando no se conoce la Varianza Poblacional.

Es frecuente que no se conozca la varianza de la población (σ2); en tales casos se debe recurrir a censos, a investigaciones similares realizadas con anterioridad o a investigaciones preliminares, denominadas encuesta piloto. Este último procedimiento es el que más se emplea para determinar el tamaño de la muestra, partiendo del supuesto de que no existe información de la población.

Es indispensable contar con un marco de referencia: la lista, mapa u otra especificación de las unidades, que resulta de la información previamente disponible, respecto a la población sobre la cual se basan los esquemas particulares de muestreo.

Posteriormente se establece la muestra, mediante:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \qquad n_0 = \frac{Z^2 S^2}{E^2} \qquad n = \frac{NZ^2 S^2}{NE^2 + Z^2 S^2}$$

Realizar la prueba piloto:

En donde $\qquad n_{preliminar}$ **= % población * N**

NOTA: Cuando n preliminar es muy pequeña es recomendable hacer un ajuste.

% población: la define el auditor, considerando costos, tamaño de la población.

Revisar archivo de Excel