Ingeniería Industrial.

Asignatura: Estadística I **Año: 3ro.** **CE.**

Clase Encuentro No. 4

Sumario.

Tema 3. Estimación.

3.1. Muestreo aleatorio simple. Propiedades.

3.2 Estimación puntual de la media, la varianza, la desviación típica y las proporciones.

3.3. Estimadores y sus propiedades.

3.4. Teorema Central del Límite. Distribución muestral de la media y la proporción.

3.5. Error Estándar de la media, Error máximo admisible y tamaño de muestra para la estimación por intervalos de la media y la proporción poblacional.

Objetivos: Que el alumno se capaz de:

- Seleccionar una Muestra Aleatorio Simple (MAS), con el empleo de apk o algún paquete de cómputos.
- Estimar los parámetros poblacionales a partir de una muestra con el empleo de apk, paquetes estadísticos y con la interpretación de las salidas que ofrecen.

Método: Expositivo.

Medios: Tiza y pizarra, Libro de texto, Tabla Estadísticas, dispositivos digitales

Forma de enseñanza: Conferencia

Bibliografía:

- Estadística para ingenieros 6ta edición R. Walpole, R. H. Myers, S.L.Myers I (1ra y 2da parte).
- Estadística Dra Caridad Bustillo y otros
- Materiales entregados y elaborados por el profesor

Situación problémica

- ✓ Se desea conocer el peso promedio que tienen las latas de conservas de tomates que se producen en una fábrica “LA CONCHITA”.



¿Qué hacer para resolver estas problemáticas?

3.1 Muestreo aleatorio simple. Propiedades.

POBLACION: Se denomina al conjunto de individuos o cosas que poseen una característica determinada. El estudio de la población se le llama **Censo**.

MUESTRA: Una parte o subconjunto de la población.

Antes de la imposibilidad material o económica de realizar un Censo, se determina tomar una muestra (MUESTREO) y a partir de esa muestra **ESTIMAR**, es decir dar un valor aproximada de los parámetros que interesa estudiar de la población.

Objeto de estudio	Medidas descriptivas	Representación
Población	Parámetros	μ, σ^2, σ, C.V.
Muestra	Estadísticos	$\overline{X}$, S^2, S, C.V.

La selección de una muestra se le llama muestreo. Existen diferentes tipos de muestreo. En el curso de trabajará con el Muestreo Aleatorio simple (MAS).

El MAS debe garantizar que todos los elementos de la población deben tener la misma probabilidad de ser seleccionados.

Existen múltiples formas de realizar un muestreo aleatorio simple (MAS). Desde una tabla estadística de números aleatorios, hasta diversos paquetes estadísticos como Excel, SPSS, aplicaciones para teléfonos móviles.

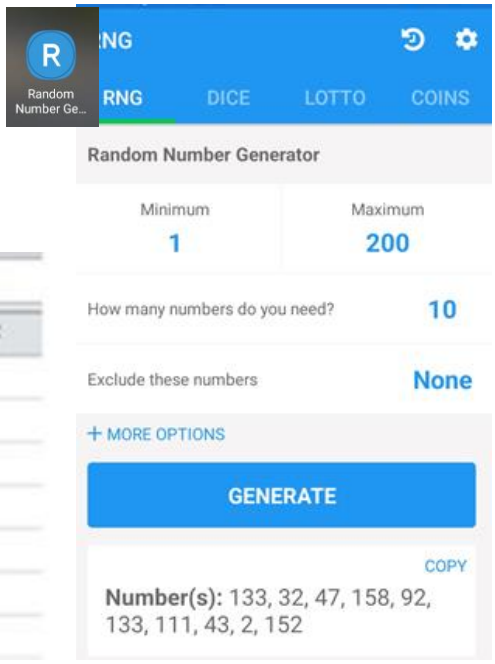
Todo comienza por asignar un número de identificación a cada valor de los elemento de la población y después generar números aleatorios para seleccionar valor del elementos identificados con ese número aleatorio

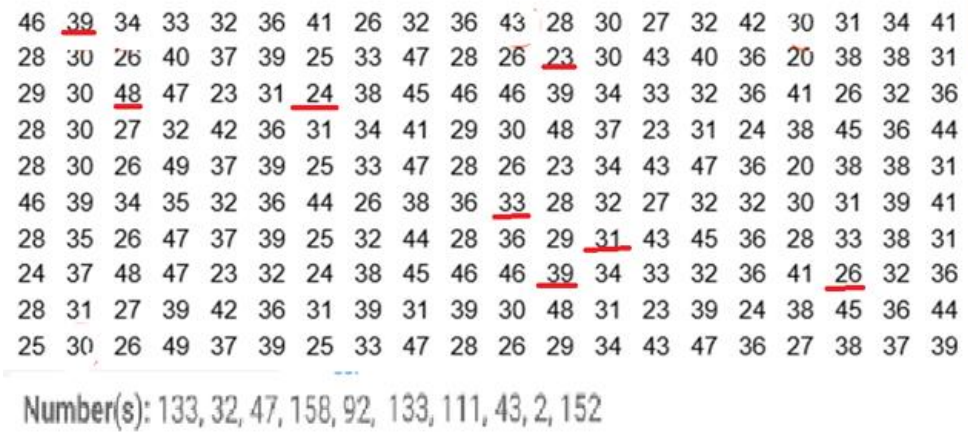
Ejemplo: Los siguientes valores corresponden a la cantidad de hojas que se perdieron por errores de impresión, durante 200 días, en un taller de impresión

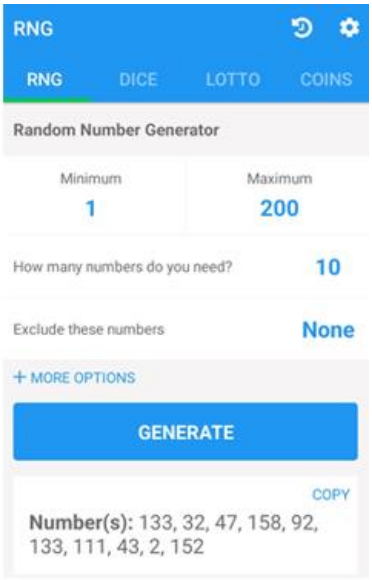
46	39	34	33	32	36	41	26	32	36	43	28	30	27	32	42	30	31	34	41
28	30	26	40	37	39	25	33	47	28	26	23	30	43	40	36	20	38	38	31
29	30	48	47	23	31	24	38	45	46	46	39	34	33	32	36	41	26	32	36
28	30	27	32	42	36	31	34	41	29	30	48	37	23	31	24	38	45	36	44
28	30	26	49	37	39	25	33	47	28	26	23	34	43	47	36	20	38	38	31
46	39	34	35	32	36	44	26	38	36	33	28	32	27	32	32	30	31	39	41
28	35	26	47	37	39	25	32	44	28	36	29	31	43	45	36	28	33	38	31
24	37	48	47	23	32	24	38	45	46	46	39	34	33	32	36	41	26	32	36
28	31	27	39	42	36	31	39	31	39	30	48	31	23	39	24	38	45	36	44
25	30	26	49	37	39	25	33	47	28	26	29	34	43	47	36	27	38	37	39

Seleccionar una muestra aleatoria al azar de las hojas que se perdieron en 5 días (n= 5)









Activar Windows

No. Aleatorio	133	32	47	158	92	133	111	43	2	152
Hojas dañadas	33	23	24	26	33	31	33	48	39	39

3.2 Estimación puntual de la media, la varianza, la desviación típica y las proporciones.

La estimación puntual consiste en considerar que estos parámetros serán igual a los mejores estadígrafos de la muestra que llamaremos (**Estimadores**).

Estimación Puntual ⇒ **μ= x̄**, **σ²= Ŝ²**, **σ= Ŝ**, **P= p̂** ⇒ **Estimadores**

Cuando se selecciona una muestra para Estimar un parámetro es muy probable que se produzca en error de muestreo que

$$e_m = \hat{\theta} - \theta \quad \text{o} \quad e_m = \bar{x} - \mu$$

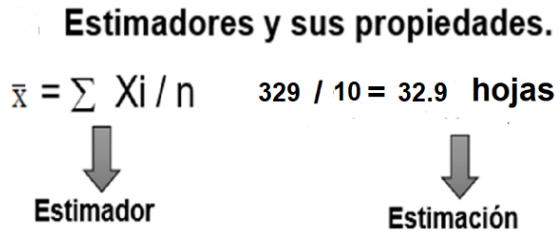
Se llamará ERROR DE MUESTREO, a la diferencia entre el valor del estimador y del parámetro. Este error de muestreo variará para cada muestra que se seleccione, por lo que es una variable aleatoria que no se puede conocer, pues habría que conocer el parámetro poblacional, y si se conociera éste, no habría necesidad de estimarlo

3.3 Estimadores y sus propiedades.

ESTIMADOR, a cualquier función de "n" variables, donde después de sustituir en ella los valores muestrales, el resultado obtenido es la ESTIMACIÓN puede servir como sustituto del valor del parámetro poblacional.

Se expresa por:

(sita circunflejo, este símbolo ^ circunflejo, denota estimación).



Parámetro	Estimador	Observaciones
μ	$\overline{X}$	
σ^2	S^2	Con el denominador (n-1)
σ	S	
P	$\hat{p}$	

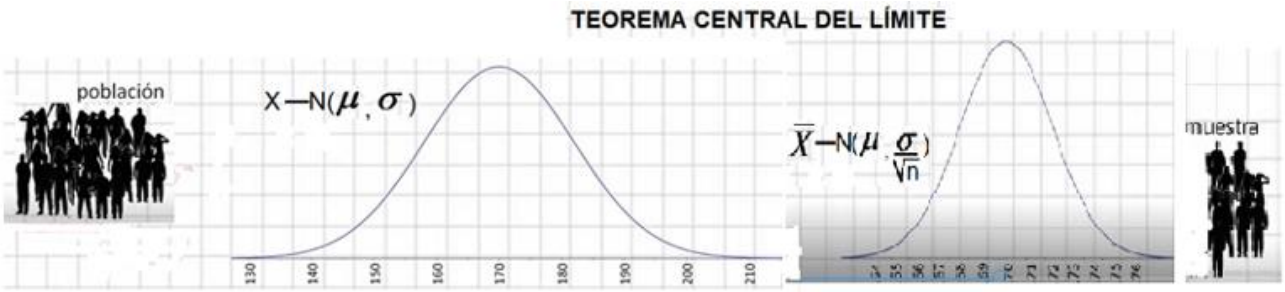
Cualidades de un buen ESTIMADOR	
INSESGADO.	$E(\hat{\theta}) = \theta$, **La esperanza del estimador es igual al parámetro.**
CONSISTENTE.	**si al hacer "n" cada vez más grande de manera que n $\rightarrow$ N, el estimador tiende a estar más cerca del parámetro**.
EFICIENTE.	si tiene varianza mínima.

3.3. Teorema Central del Límite. Distribución muestral de la media y la proporción.

Distribución muestral de la muestra

Si se hace una selección de todas las muestras posibles, a la distribución de los resultados de cada una de las muestras se llama DISTRIBUCIÓN MUESTRAL.

Teorema Central del Límite:



Si una población tiene distribución NORMAL con media μ y desviación estándar σ conocida, y tomamos muestras de cualquier tamaño.

Entonces la distribución de la media muestral sigue una distribución Normal con media μ y Desviación estándar σ / √n

Para la Estandarización de la variable $Z = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

y para calcular la probabilidad de cierto comportamiento de la media, se utilizará la variable estandarizada:

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

Si la variable X de la población sigue una distribución Normal se puede conocer o no conocer su varianza o desviación típica

ESTIMADOR	POBLACIÓN CON DISTRIBUCIÓN NORMAL (N)	DISTRIBUCIÓN MUESTRA
MEDIA	Se conoce la desviación típica de la población	si $X \rightarrow N(\mu, \sigma) \Rightarrow$ $\bar{X} \rightarrow N(\mu, \sigma/\sqrt{n})$
	No se conoce la desviación típica de la población	si $X \rightarrow N(\mu, ?) \Rightarrow$ $\bar{X} \rightarrow N(\mu, s/\sqrt{n})$

Si la variable X de la población no sabemos su distribución, con media μ y desviación estándar desconocida

ESTIMADOR	NO SE CONOCE LA DISTRIBUCIÓN DE LA POBLACIÓN	DISTRIBUCIÓN MUESTRA
MEDIA	No se conoce la desviación típica de la población y la muestra es mayor que 30)	si $X \rightarrow ?(\mu, ?)$ y $n \geq 30 \Rightarrow$ $\bar{X} \approx N(\mu, S/\sqrt{n})$
	No se conoce la desviación típica de la población y la muestra es menor que 30)	si $X \rightarrow ?(\mu \ ?)$ y $n < 30 \Rightarrow$ $\bar{X} \rightarrow t_{(n-1;\ 1-\alpha/2)}(\mu, s/\sqrt{n})$

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

EN LA PRACTICA SE HA DEMOSTRADO QUE SIEMPRE QUE N ≥ 30 LA APROXIMACION A LA NORMAL ES BUENA, POR LO QUE SE UTILIZARA E ESTE CRITERIO PARA CONSIDERAR QUE n → ∞

Distribución muestral de proporciones.

PROPORCIÓN	**Se plantea que esta aproximación es realmente buena cuando np > 5 ó nq < 5**	**x/n → N (P, $\sqrt{pq/n}$) al menos aproximadamente**

x/n → N (P, $\sqrt{pq/n}$) al menos aproximadamente. Se plantea que esta aproximació realmente buena cuando np > 5 y/o nq < 5.

x/n → N (P, $\sqrt{pq/n}$) al menos aproximadamente. Se plantea que esta aproximación es realmente buena cuando np > 5 y/o nq < 5.

3.5. Error Estándar de la media, Error máximo admisible y tamaño de muestra para la estimación por intervalos de la media y la proporción poblacional.

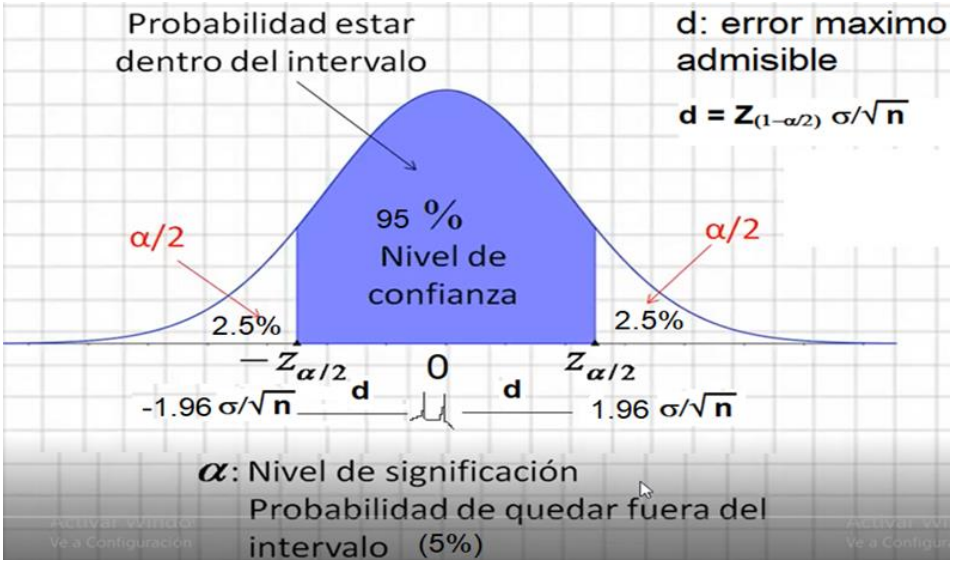
Error Estándar (Típico) de la medias (E_S)

$$Es = \frac{\sigma}{\sqrt{n}}$$

Este dato indica qué tanto la media o el promedio de la muestra se acerca al promedio real de la población de donde se extrajo el conjunto de datos. Mientras más larga sea la muestra, más pequeño será el error estándar, y más próximo estará el promedio de la muestra al promedio de la población.

Error máximo admisible (d):

$$d = Z_{(1-\alpha/2)}\ \sigma/\sqrt{n}$$



El producto del coeficiente de confianza $Z_{(1-\alpha/2)}$ y el Error Estándar del estimador

$$d = Z_{(1-\alpha/2)}\ \sigma/\sqrt{n}$$

Intervalo de confianza para la media

$$IC\ (\mu) = (\bar{x} \pm d)$$

$$IC\ (\mu) = (\bar{x} - d \leq \mu \leq \bar{x} + d)$$

Intervalo de confianza (IC) para la proporción.

$$d = Z_{(1-\alpha/2)}\ \sqrt{pq/n}$$

$$IC = (p \pm d)$$

$$(p - d \leq P \leq p + d)$$

Conclusiones

ESTIMADOR	POBLACIÓN CON DISTRIBUCIÓN NORMAL (N)	DISTRIBUCIÓN MUESTRA	" d " y " n "
MEDIA	Se conoce la desviación típica de la población	si $X \to N(\mu, \sigma) \Rightarrow \bar{X} \to N(\mu, \sigma/\sqrt{n})$	$d = Z_{(1-\alpha/2)} * \sigma / \sqrt{n}$ $n = [Z_{(1-\alpha/2)} * \sigma / d]^2$
	No se conoce la desviación típica de la población	si $X \to N(\mu, ?) \Rightarrow \bar{X} \to N(\mu, s/\sqrt{n})$	$d = Z_{(1-\alpha/2)} * s / \sqrt{n}$ $n = [Z_{(1-\alpha/2)} * s / d]^2$
ESTIMADOR	**NO SE CONOCE LA DISTRIBUCIÓN DE LA POBLACIÓN**	**DISTRIBUCIÓN MUESTRA**	**" d " y " n "**
MEDIA	No se conoce la desviación típica de la población y la muestra es mayor que 30)	si $X \to ?(\mu, ?)$ y $n \geq 30 \Rightarrow \bar{X} \approx N(\mu, S/\sqrt{n})$	$d = Z_{(1-\alpha/2)} * S / \sqrt{n}$ $n = [Z_{(1-\alpha/2)} * S / d]^2$
	No se conoce la desviación típica de la población y la muestra es menor que 30)	si $X \to ?(\mu \ ?)$ y $n < 30 \Rightarrow \bar{X} \to t_{(n-1;\ 1-\alpha/2)} (\mu, s/\sqrt{n})$	$d = t_{(n-1;1-\alpha/2)} * S / \sqrt{n})$ $n = [t_{(n-1;1-\alpha/2)} * S / d]^2$
PROPORCIÓN	Se plantea que esta aproximación es realmente buena cuando $np > 5$ ó $nq < 5$	$x/n \to N(P, \sqrt{pq/n})$ al menos aproximadamente	$d = Z_{(1-\alpha/2)} * \sqrt{pq/n}$ $n = [Z^2_{(1-\alpha/2)} * pq]/d^2$

ESTUDIO INDEPENDIENTE.

Los siguientes valores corresponden a los ingresos diarios en CUP de todos los trabajadores de una empresa de la Artemisa.

46	39	34	33	32	36	41	26	32	36	43	28	30	27	32	42	30	31	34	41
28	30	26	40	37	39	25	33	47	28	26	23	30	43	40	36	20	38	38	31
29	30	48	47	23	31	24	38	45	46	46	39	34	33	32	36	41	26	32	36
28	30	27	32	42	36	31	34	41	29	30	48	37	23	31	24	38	45	36	44
28	30	26	49	37	39	25	33	47	28	26	23	34	43	47	36	20	38	38	31

a) Seleccione una muestra aleatoria de 10 trabajadores y estime el promedio de los ingresos diarios.
b) A partir de la muestra seleccionada que proporción de trabajadores tienen ingresos superiores a 35 CUP.